(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 November 2002 (21.11.2002)

PCT

(10) International Publication Number
WO 02/092015 A2

(51) International Patent Classification[7]: A61K

(21) International Application Number: PCT/US02/15982

(22) International Filing Date: 17 May 2002 (17.05.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/291,311	17 May 2001 (17.05.2001)	US
60/353,058	1 February 2002 (01.02.2002)	US
60/361,293	4 March 2002 (04.03.2002)	US

(71) Applicants (*for all designated States except US*): **GENOME THERAPEUTICS CORPORATION** [US/US]; 100 Beaver Street, Waltham, MA 02453 (US). **WYETH** [US/US]; Five Giralda Farms, Madison, NJ 07928 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): **ALLEN, Kristina** [US/US]; 11 Oliver Lane, Hopkinton, MA 01748-3108 (US). **ANISOWICZ, Anthony** [US/US]; 50 Upham Street, West Newton, MA 02465 (US). **BHAT, Bheem, M.** [IN/US]; 1214 Mayapple Lane, West Chester, PA 19380 (US). **DAMAGNEZ, Veronique** [FR/US]; 125 Water Street, Framingham, MA 01701 (US). **ROBINSON, John, Allen** [US/US]; 23 Webb Road, Downingtown, PA 19335 (US). **YAWORSKY, Paul, J.** [US/US]; 13 Hobart Lane, Rockland, MA 02370 (US).

(74) Agents: **REA, Teresa, Stanek** et al.; Burns, Doane, Swecker & Mathis L.L.P., P.O. Box 1404, Alexandria, VA 22313-1404 (US).

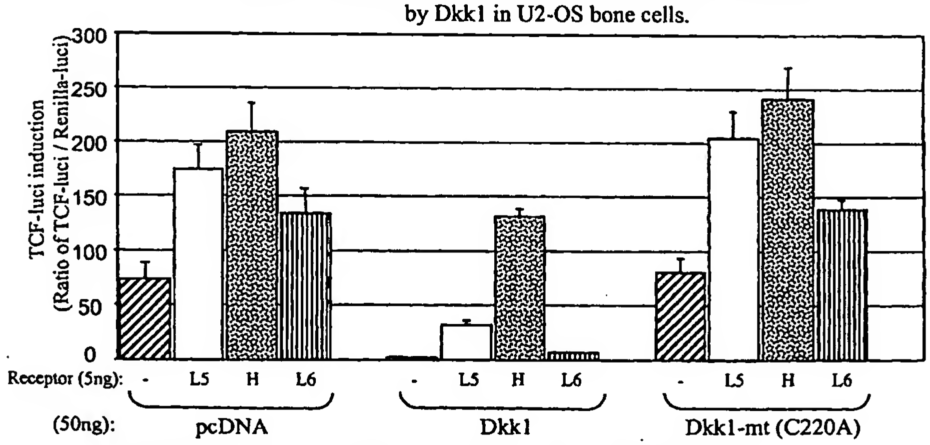
(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),

[Continued on next page]

(54) Title: REAGENTS AND METHODS FOR MODULATING DKK-MEDIATED INTERACTIONS

Wnt1 - HBM generated TCF-luci is not efficiently inhibited by Dkk1 in U2-OS bone cells.



• With Wnt1 the TCF-signal generated by LRP5 is greater than that of LRP6.
• LRP5/6 -Wnt1 induced TCF- is efficiently blocked byDkk1

(57) Abstract: The present invention provides reagents, compounds, compositions, and methods relating to novel interactions of the extracellular domain of LRP5, HBM (a variant of LRP5), and/or LRP6 with Dkk, including Dkk-1. The various nucleic acids, polypeptides, antibodies, assay methods, diagnostic methods, and methods of treatment of the present invention are related to and impact on Dkk, LRP5, LRP6, HBM, and Wnt signaling. Dkk, LRP5, LRP6, HBM, and Wnt are implicated in bone and lipid cellular signaling. Thus, the present invention provides reagents and methods for modulating lipid levels and/or bone mass and is useful in the treatment and diagnosis of abnormal lipid levels and bone mass disorders, such as osteoporosis.

WO 02/092015 A2

Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

REAGENTS AND METHODS FOR MODULATING DKK-MEDIATED INTERACTIONS

FIELD OF THE INVENTION

The present invention relates to signal transduction, bone development, bone loss disorders, modulation of lipid-related conditions, research reagents, methods of screening drug leads, drug development, treatments for bone and/or lipid disorders, screening and development of therapies, molecular, cellular, and animal models of bone and/or lipid development and maintenance, which are mediated by Dkk, including Dkk-1, and/or LRP5, LRP6, HBM or other members of the Wnt pathway.

BACKGROUND OF THE INVENTION

Two of the most common types of osteoporosis are postmenopausal and senile osteoporosis. Osteoporosis affects both men and women, and, taken with other abnormalities of bone, presents an ever-increasing health risk for an aging population. The most common type of osteoporosis is that associated with menopause. Most women lose between 20-60% of the bone mass in the trabecular compartment of the bone within 3-6 years after the cessation of menses. This rapid bone loss is generally associated with an increase of bone resorption and formation. However, the resorptive cycle is more dominant and the result is a net loss of bone mass. Osteoporosis is a common and serious disease among postmenopausal women. There are an estimated 25 million women in the United States alone who are afflicted with this disease. The results of osteoporosis are personally harmful, and also account for a large economic loss due to its chronicity and the need for extensive and long-term support (*e.g.,* hospitalization and nursing home care) from disease sequelae. This is especially true in elderly patients. Additionally, while osteoporosis is generally not thought of as a life-threatening condition, a 20-30% mortality rate is related to hip fractures in elderly women. A large percentage of this mortality rate can be directly associated with postmenopausal osteoporosis.

The most vulnerable tissue in the bone to the effects of postmenopausal osteoporosis is the trabecular bone. This tissue is often referred to as spongy bone and is particularly concentrated near the ends of the bone, near the joints, and in the vertebrae of the spine. The trabecular tissue is characterized by small structures which inter-connect with each other as well as the more solid and dense cortical tissue which makes up the outer surface and central shaft of the bone. This criss-cross network of trabeculae gives lateral support to the outer cortical structure and is critical to the biomechanical strength of the overall structure. In postmenopausal osteoporosis, it is primarily the net resorption and loss of the trabeculae which lead to the failure and fracture of the bone. In light of the loss of the trabeculae in postmenopausal women, it is not surprising that the most common fractures are those associated with bones which are highly dependent on trabecular support, *e.g.*, the vertebrae, the neck of the femur, and the forearm. Indeed, hip fracture, Colle's fractures, and vertebral crush fractures are indicative of postmenopausal osteoporosis. Osteoporosis affects cortical as well as trabecular bone. Alterations in endosteal bone resorption and Haversian remodeling with age affect cortical thickness and structural integrity contributing the increased risk for fracture.

One of the earliest generally accepted methods for treatment of postmenopausal osteoporosis was estrogen replacement therapy. Although this therapy frequently is successful, patient compliance is low, primarily due to the undesirable side-effects of chronic estrogen treatment. Frequently cited side-effects of estrogen replacement therapy include reinitiation of menses, bloating, depression, and, potentially, increased risk of breast or uterine cancer. In order to limit the known threat of uterine cancer in women who have not had a hysterectomy, a protocol of estrogen and progestin cyclic therapy is often employed. This protocol is similar to that used in birth control regimens, and often is not tolerated by women because of the side-effects characteristic of progestin. More recently, certain antiestrogens, originally developed for the treatment of breast cancer, have been shown in experimental models of postmenopausal osteoporosis to be efficacious. Among these agents is raloxifene (*See*, U.S. Patent No. 5,393,763; Black *et al.*, *J.*

Clin. Invest., 93:63-69 (1994); and Ettinger *et al.*, *JAMA* 282:637-45 (1999)). In addition, tamoxifen, a widely used clinical agent for treating breast cancer, has been shown to increase bone mineral density in post menopausal women suffering from breast cancer (Love *et al.*, *N. Engl. J. Med.*, 326:852-856 (1992)).

Another therapy for the treatment of postmenopausal osteoporosis is the use of calcitonin. Calcitonin is a naturally occurring peptide which inhibits bone resorption and has been approved for this use in many countries (Overgaard *et al.*, *Br. Med. J.*, 305:556-561 (1992)). The use of calcitonin has been somewhat limited, however. Its effects are very modest in increasing bone mineral density, and the treatment is very expensive. Another therapy for the treatment of postmenopausal osteoporosis is the use of bisphosphonates. These compounds were originally developed for treating Paget's disease and malignant hypercalcemia. They have been shown to inhibit bone resorption. Alendronate, a bisphosphonate, has been approved for the treatment of postmenopausal osteoporosis. These agents may be helpful in the treatment of osteoporosis, but these agents also have potential liabilities which include osteomalacia, extremely long half-life in bone (greater than 2 years), and possible "frozen bone syndrome," *e.g.*, the cessation of normal bone remodeling.

Senile osteoporosis is similar to postmenopausal osteoporosis in that it is marked by the loss of bone mineral density and resulting increase in fracture rate, morbidity, and associated mortality. Generally, it occurs in later life, *i.e.*, after 70 years of age. Historically, senile osteoporosis has been more common in females, but with the advent of a more elderly male population, this disease is becoming a major factor in the health of both sexes. It is not clear what, if any, role hormones such as testosterone or estrogen have in this disease, and its etiology remains obscure. Treatment of this disease has not been very satisfactory. Hormone therapy, estrogen in women and testosterone in men, has shown equivocal results; calcitonin and bisphosphonates may be of some utility.

The peak mass of the skeleton at maturity is largely under genetic control. Twin studies have shown that the variance in bone mass between adult monozygotic

twins is smaller than between dizygotic twins (Slemenda *et al.*, *J. Bone Miner. Res.*, 6: 561-567 (1991); Young *et al.*, *J. Bone Miner. Res.*, 6:561-567 (1995); Pocock *et al.*, *J. Clin. Invest.*, 80:706-710 (1987); Kelly *et al.*, *J. Bone Miner. Res.*, 8:11-17 (1993)). It has been estimated that up to 60% or more of the variance in skeletal mass is inherited (Krall *et al.*, *J. Bone Miner. Res.*, 10:S367 (1993)). Peak skeletal mass is the most powerful determinant of bone mass in elderly years (Hui *et al.*, *Ann. Int. Med.*, 111:355-361 (1989)), even though the rate of age-related bone loss in adult and later life is also a strong determinant (Hui *et al.*, *Osteoporosis Int.*, 1:30-34 (1995)). Since bone mass is the principal measurable determinant of fracture risk, the inherited peak skeletal mass achieved at maturity is an important determinant of an individual's risk of fracture later in life. Thus, study of the genetic basis of bone mass is of considerable interest in the etiology of fractures due to osteoporosis.

Recently, a strong interest in the genetic control of peak bone mass has developed in the field of osteoporosis. The interest has focused mainly on candidate genes with suitable polymorphisms to test for association with variation in bone mass within the normal range, or has focused on examination of genes and gene loci associated with low bone mass in the range found in patients with osteoporosis. The vitamin D receptor locus (VDR) (Morrison *et al.*, *Nature*, 367:284-287 (1994)), PTH gene (Howard *et al.*, *J. Clin. Endocrinol. Metab.*, 80:2800-2805 (1995); Johnson *et al.*, *J. Bone Miner. Res.*, 8:11-17 (1995); Gong *et al.*, *J. Bone Miner. Res.*, 10:S462 (1995)) and the estrogen receptor gene (Hosoi *et al.*, *J. Bone Miner. Res.*, 10:S170 (1995); Morrison *et al.*, *Nature*, 367:284-287 (1994)) have figured most prominently in this work. These studies are difficult because bone mass (*i.e*, the phenotype) is a continuous, quantitative, polygenic trait, and is confounded by environmental factors such as nutrition, co-morbid disease, age, physical activity, and other factors. Also, this type of study design requires large numbers of subjects. In particular, the results of VDR studies to date have been confusing and contradictory (Garnero *et al.*, *J. Bone Miner. Res.*, 10:1283-1288 (1995); Eisman *et al.*, *J. Bone. Miner. Res.*, 10:1289-1293 (1995); Peacock, *J. Bone Miner. Res.*, 10:1294-1297 (1995)).

Furthermore, thus far, the art has not determined the mechanism(s) whereby the genetic influences exert their effect on bone mass.

While it is well known that peak bone mass is largely determined by genetic rather than environmental factors, studies to determine the gene loci (and ultimately the genes) linked to variation in bone mass are difficult and expensive. Study designs which utilize the power of linkage analysis, *e.g.*, sib-pair or extended family, are generally more informative than simple association studies, although the latter do have value. However, genetic linkage studies involving bone mass are hampered by two major problems. The first problem is the phenotype, as discussed briefly above. Bone mass is a continuous, quantitative trait, and establishing a discrete phenotype is difficult. Each anatomical site for measurement may be influenced by several genes, many of which may be different from site to site. The second problem is the age component of the phenotype. By the time an individual can be identified as having low bone mass, there is a high probability that their parents or other members of prior generations will be deceased and therefore unavailable for study, and younger generations may not have even reached peak bone mass, making their phenotyping uncertain for genetic analysis.

Thus, there is a need in the art for additional research tools for the elucidation of the molecular mechanism of bone modulation, for the screening and development of candidate drugs, and for treatments of bone development and bone loss disorders. The present invention is directed to these, as well as other, important ends.

In addition to bone modulation, the present invention relates to modulation of lipid levels. Cardiovascular disease is the most common cause of mortality in the United States, and atherosclerosis is the major cause of heart disease and stroke. It is widely appreciated that cholesterol plays an important role in atherogenesis. Normally, most cholesterol serves as a structural element in the walls of cells, whereas much of the rest is in transit through the blood or functions as the starting material for the synthesis of bile acids in the liver, steroid hormones in endocrine cells and vitamin D in skin. The transport of cholesterol and other lipids through the

circulatory system is facilitated by their packaging into lipoprotein carriers. These spherical particles comprise protein and phospholipid shells surrounding a core of neutral lipid, including unesterified ("free") or esterified cholesterol and triglycerides. Risk for atherosclerosis increases with increasing concentrations of low density lipoprotein (LDL) cholesterol, whereas risk is inversely proportional to levels of high-density lipoprotein (HDL) cholesterol. The receptor-mediated control of plasma LDL levels has been well-defined, and recent studies have now provided new insights into HDL metabolism.

The elucidation of LDL metabolism began in 1974 by Michael Brown and Joseph Goldstein. In brief, the liver synthesizes a precursor lipoprotein (very low density lipoprotein, VLDL) that is converted during circulation to intermediate density lipoprotein (IDL) and then to LDL. The majority of the LDL receptors expressed in the body are on the surfaces of liver cells, although virtually all other tissues ("peripheral tissues") express some LDL receptors. After binding, the receptor-lipoprotein complex is internalized by the cells via coated pits and vesicles, and the entire LDL particle is delivered to lysosomes, wherein it is dissembled by enzymatic hydrolysis, releasing cholesterol for subsequent cellular metabolism. This whole-particle uptake pathway is called "receptor-mediated endocytosis." Cholesterol-mediated feedback regulation of both the levels of LDL receptors and cellular cholesterol biosynthesis help ensure cellular cholesterol homeostasis. Genetic defects in the LDL receptor in humans results in familial hypercholesterolemia, a disease characterized by elevated plasma LDL cholesterol and premature atherosclerosis and heart attacks. One hypothesis for the deleterious effects of excess plasma LDL cholesterol is that LDL enters the artery wall, is chemically modified, and then is recognized by a special class of receptors called macrophage scavenger receptors, that mediate the cellular accumulation of the LDL cholesterol in the artery, eventually leading to the formation of an atherosclerotic lesion.

The major lipoprotein classes include intestinally derived chylomicrons that transport dietary fats and cholesterol, hepatic-derived VLDL, IDL, and LDL that can be atherogenic, and hepatic- and intestinally-derived HDL that are antiatherogenic.

Apoprotein B (ApoB) is necessary for the secretion of chylomicrons (ApoB48) and VLDL, IDL, and LDL (ApoB100). Plasma levels of VLDL triglycerides are determined mainly by the rates of secretion in LDL lipolytic activity. Plasma levels of LDL cholesterol are determined mainly by the secretion of ApoB100 into plasma, the efficacy with which VLDL are converted to LDL and by LDL receptor-mediated clearance. Regulation of HDL cholesterol levels is complex and is affected by rates of synthesis of its Apo proteins, rates of esterification of free cholesterol to cholesterol ester by LCAT, levels of triglyceride-rich lipoproteins and CETP-mediated transfer of cholesterol esters from HDL, and clearance from plasma of HDL lipids and Apo proteins.

Normal lipoprotein transport is associated with low levels of triglycerides and LDL cholesterol and high levels of HDL cholesterol. When lipoprotein transport is abnormal, lipoprotein levels can change in ways that predispose individuals to atherosclerosis and arteriosclerosis (see Ginsburg, *Endocrinol. Metab. Clin. North Am.*, 27:503-19 (1998)).

Several lipoprotein receptors may be involved in cellular lipid uptake. These receptors include: scavenger receptors; LDL receptor-related protein/α2-macroglobulin receptor (LRP); LDL receptor; and VLDL receptor. With the exception of the LDL receptor, all of these receptors are expressed in atherosclerotic lesions while scavenger receptors are mostly expressed in macrophages, the LRP and VLDL receptors may play an important role in mediating lipid uptake in smooth muscle cells (Hiltunen *et al.*, *Atherosclerosis*, 137 suppl.:S81-8 (1998)).

A major breakthrough in the pharmacologic treatment of hypercholesterolemia has been the development of the "statin" class of 3-hydroxy-3-methylglutaryl-CoA reductase (HMG CoA reductase) inhibitory drugs. 3-hydroxy-3-methylglutaryl-CoA reductase is the rate controlling enzyme in cholesterol biosynthesis, and its inhibition in the liver stimulates LDL receptor expression. As a consequence, both plasma LDL cholesterol levels and the risk for atherosclerosis decrease. The discovery and analysis of the LDL receptor system has had a profound impact on cell biology, physiology, and medicine.

HDL is thought to remove unesterified, or "free" cholesterol (FC) from peripheral tissues, after which most of the cholesterol is converted to cholesterol ester (CE) by enzymes in the plasma. Subsequently, HDL cholesterol is efficiently delivered directly to the liver and steroidogenic tissues via a selective uptake pathway and the HDL receptor, SR-BI (class B type I scavenger receptor) or, in some species, transferred to other lipoproteins for additional transport in metabolism (see Krieger, *Proc. Natl. Acad. Sci. USA*, 95:4077-4080 (1998)).

These issues illustrate a need in the art for additional research tools for the elucidation of the molecular mechanism of lipid modulation, for the screening and development of candidate drugs, and for treatments of lipid levels and lipid level modulation disorders. The present invention is directed to these, as well as other, important ends.

SUMMARY OF THE INVENTION

The present invention provides reagents, compounds, compositions and methods relating to novel interactions of the extracellular domain of LRP5, HBM (a variant of LRP5), and/or LRP6 with Dkk proteins. LRP5 is also referred to as Zmax1 or Zmax. Thus, when discussing methods, reagents, compounds, and compositions of the invention which relate to the interaction between Dkk and LRP5 (or Zmax1), the invention is also to be understood to encompass embodiments relating to interactions between Dkk and LRP6 and Dkk and HBM. Moreover, where Dkk is discussed herein, it is to be understood that the methods, reagents, compounds, and compositions of the present invention include the Dkk family members, including but not limited to Dkk-1, Dkk-2, Dkk-3, Dkk-4 and Soggy. Furthermore, the invention encompasses novel fragments of Dkk-1 which demonstrate a binding interaction between the ligand binding domain (LBD) of LRP5 and additional proteins and/or which can modulate an interaction between LRP5, or a variant or fragment thereof, and a Dkk protein. The invention provides assays, methods, compositions, and compounds relating to Dkk-Wnt signaling. Numerous Wnt proteins are compatible with the present invention, including Wnt1-Wnt19, and particularly, Wnt1, Wnt3,

Wnt3a, and Wnt10b. The present invention further provides reagents, compounds, compositions and methods modulating interactions between one or more other proteins and Dkk-1. The present invention also provides a series of peptide aptamers which bind to Dkk-1 or to LRP5 (or HBM and/or LRP6).

The polypeptides of the invention, for example in the form of peptide oligomers, aptamers, proteins, and protein fragments as well as the nucleic acids of the invention, for example in the form of nucleic acids which encode the polypeptides of the invention as well as antisense, or complimentary nucleic acids, are useful as reagents for the study of bone mass and lipid level modulation. The polypeptides and nucleic acids of the invention are also useful as therapeutic and diagnostic agents.

The present invention provides useful reagents for the modulation of Dkk proteins with LRP5, LRP6, and/or HBM, the modulation Dkk-1 and/or Dkk-1 interacting protein activity, and modulation of LRP5/Dkk-1, LRP6/Dkk1 and HBM/Dkk-1 interactions and Dkk-1/Dkk-1 interacting protein interactions. The present invention provides a series of peptide aptamers which bind Dkk-1 or LRP5, LRP6, and/or HBM.

An object of the invention is to provide for a method of regulating LRP5/LRP6/HBM/HBM-like activity in a subject comprising administering a therapeutically effective amount of a composition which modulates Dkk activity. The subject can be a vertebrate or an invertebrate organism, but more preferably the organism is a canine, a feline, an ovine, a primate, an equine, a porcine, a caprine, a camelid, an avian, a bovine, or a rodent organism. A more preferred organism is a human. In a preferred embodiment, the Dkk protein is Dkk-1. In a particularly preferred embodiment, Dkk-1 activity is decreased. In another embodiment, Dkk activity modulates bone mass and/or lipid levels. In a preferred embodiment, bone mass is increased and/or lipid levels are decreased. In another preferred embodiment, the modulation in bone mass is an increase in bone strength determined via one or more of a decrease in fracture rate, an increase in areal bone density, an increase in volumetric mineral bone density, an increase in trabecular

connectivity, an increase in trabecular density, an increase in cortical density or thickness, an increase in bone diameter, and an increase in inorganic bone content. The invention further provides such a method wherein the composition comprises a Dkk, Dkk-1 or a LRP5/LRP6/HBM binding fragment thereof, such as those depicted in Figure 6 or a mimetic of those fragments depicted in Figure 6. The invention further provides such a method wherein the composition comprises one or more of the proteins which interact with Dkk, including Dkk-1, such as those depicted in Figure 5, or a Dkk-binding fragment thereof, or an antisense, siRNA, or shRNA molecule which recognizes and binds to a nucleic acid encoding one or more Dkk interacting or Dkk-1 interacting proteins. The invention further provides such a method wherein the composition comprises an LRP5/LRP6/Zmax1 antibody, Dkk antibody, a Dkk-1 antibody or an antibody to a Dkk-1 interacting protein. The invention further provides such a method wherein the compositions comprise an aptamer of Dkk or Dkk-1, such as those depicted in Figure 3 (SEQ ID NOs:171-188), or a mimetic of such an aptamer. The method further provides that invention further provides such a method wherein the compositions comprise an aptamer of a Dkk interacting or Dkk-1 interacting protein, or a mimetic of such an aptamer.

A composition of the present invention may modulate activity either by enhancing or inhibiting the binding of Dkk to LRP5/LRP6/Zmax1, particularly Dkk-1, or the binding of Dkk-1 to a Dkk-1 interacting protein, such as those shown in Figure 5. A composition of the present invention may comprise an LRP5 peptide aptamer, such as OST262 (SEQ ID NO:208), Figures 4 (SEQ ID NOs:189-192) (particularly, peptide (SEQ ID NO:191) and 13 (including SEQ IDNOs:204-214), or a mimetic of such an aptamer. Preferred compositions of the present invention also comprise LRP5 antibodies.

Another aspect of the invention is to provide for a method of regulating Dkk-Wnt pathway activity in a subject comprising administering a therapeutically effective amount of a composition which modulates Dkk-Wnt pathway activity. In a preferred embodiment, the Dkk protein is Dkk-1. In a particularly preferred embodiment, Dkk-1 activity is decreased. In another embodiment, Dkk activity modulates bone mass

and/or lipid levels. In a preferred embodiment, bone mass is increased and/or lipid levels are decreased. In another preferred embodiment, the modulation in bone mass is an increase in bone strength determined via one or more of a decrease in fracture rate, an increase in areal bone density, an increase in volumetric mineral bone density, an increase in trabecular connectivity, an increase in trabecular density, an increase in cortical density or thickness, an increase in bone diameter, and an increase in inorganic bone content. In another preferred embodiment, the Wnt is Wnt1-Wnt19. In a particularly preferred embodiment, the Wnt is Wnt1, Wn3, Wnt3a, or Wnt10b. Preferred compositions comprise Dkk-modulating or Dkk-1-modulating compounds or one or more Dkk interacting or Dkk-1 interacting proteins, or a Dkk-binding fragment thereof. Other preferred Dkk modulating compositions comprise a Dkk or Dkk-1 antibody or an antibody to a Dkk interacting or Dkk-1 interacting protein. Also contemplated are antisense, siRNA, and shRNA molecules which recognize and bind to a nucleic acid encoding one or more Dkk-1 interacting proteins. The invention further provides such a method wherein the composition comprises a biologically active or LRP5/LRP6/HBM binding fragment of Dkk, including Dkk-1, such as those depicted in Figure 6 or a mimetic of those fragments depicted in Figure 6. The Dkk modulating composition may also comprise a peptide aptamer of a Dkk interacting or Dkk-1 interacting protein, or a mimetic of such an aptamer. A composition of the present invention may modulate activity either by enhancing or inhibiting the binding of Dkk, including Dkk-1, to LRP5, LRP6, or HBM or the binding of Dkk, including Dkk-1, to a Dkk interacting protein, such as those shown in Figure 5. The invention further provides such a method wherein the composition comprises an aptamer of Dkk or Dkk-1, such as those depicted. A composition of the present invention may comprise an LRP5 peptide aptamer, such as OST262 (SEQ ID NO:208). Preferred compositions of the present invention also comprise LRP5 antibodies.

A further aspect of the invention is to provide for a method of modulating Wnt signaling in a subject comprising administering a therapeutically effective amount of a composition which modulates Dkk activity or modulates Dkk interaction with LRP5

(or LRP6 or HBM). In a preferred embodiment, the Dkk protein is Dkk-1. In a particularly preferred embodiment, Dkk-1 activity is decreased. In another embodiment, Dkk activity modulates bone mass and/or lipid levels. In a preferred embodiment, bone mass is increased and/or lipid levels are decreased. In another preferred embodiment, the modulation in bone mass is an increase in bone strength determined via one or more of a decrease in fracture rate, an increase in areal bone density, an increase in volumetric mineral bone density, an increase in trabecular connectivity, an increase in trabecular density, an increase in cortical density or thickness, an increase in bone diameter, and an increase in inorganic bone content. In another preferred embodiment, the Wnt is Wnt1-Wnt19. In a particularly preferred embodiment, the Wnt is Wnt1, Wnt3, Wnt3a, or Wnt10b. Preferred Wnt modulating compositions comprise one or more Dkk interacting or Dkk-1 interacting proteins, or a biologically active or LRP5/LRP6/HBM binding fragment thereof. Also contemplated are antisense, siRNA, and shRNA molecules which recognize and bind to a nucleic acid encoding one or more Dkk interacting or Dkk-1 interacting proteins. The invention further provides such a method wherein the composition comprises a biologically active or LRP5/LRP6/HBM binding fragment of Dkk or Dkk-1, such as those depicted in Figure 6 or a mimetic of those fragments depicted in Figure 6. The Dkk modulating composition may also comprise a peptide aptamer of a Dkk interacting or Dkk-1 interacting protein, or a mimetic of such an aptamer. A composition of the present invention may modulate activity either by enhancing or blocking the binding of Dkk, including Dkk-1, to LRP5, LRP6, or HBM or the binding of Dkk or Dkk-1 to a Dkk interacting or Dkk-1 interacting protein, such as those shown in Figure 5. The invention further provides such a method wherein compositions comprising an aptamer of Dkk or Dkk-1, such as those depicted in Figure 3 (SEQ ID NOs:171-188), or a mimetic of such an aptamer. The invention further provides such a method wherein the composition comprises a Dkk or Dkk-1 antibody or an antibody to a Dkk interacting or Dkk-1 interacting protein. The invention further provides such a method wherein compositions of an LRP5 peptide aptamer, such as OST262 (SEQ ID NO:208), Figures 4 (SEQ ID NO:189-192

(particularly peptide (SEQ ID NO:191) and Figure 13 (including SEQ ID NOs:204-214), or a mimetic of such an aptamer. Additional preferred compositions of the present invention also comprise LRP5 antibodies.

Additionally, the invention provides for a method of modulating bone mass and/or lipid levels in a subject comprising administering to the subject a composition which modulates Dkk activity or Dkk interaction with LRP5 in an amount effective to modulate bone mass and/or lipid levels, wherein bone mass and/or lipid levels are in need of modulation. In a preferred embodiment, the Dkk protein is Dkk-1. In a particularly preferred embodiment, Dkk-1 activity is decreased. In another embodiment, Dkk activity modulates bone mass and/or lipid levels. In a preferred embodiment, bone mass is increased and/or lipid levels are decreased. In another preferred embodiment, the modulation in bone mass is an increase in bone strength determined via one or more of a decrease in fracture rate, an increase in areal bone density, an increase in volumetric mineral bone density, an increase in trabecular connectivity, an increase in trabecular density, an increase in cortical density or thickness, an increase in bone diameter, and an increase in inorganic bone content. Preferred bone mass and/or lipid modulating compositions comprise one or more Dkk interacting or Dkk-1 interacting proteins, or a biologically active or LRP5/LRP6/HBM binding fragment thereof. Also contemplated are antisense, siRNA, and shRNA molecules which recognize and bind to a nucleic acid encoding one or more Dkk interacting or Dkk-1 interacting proteins. The invention further provides such a method wherein the composition comprises a biologically active or LRP5/LRP6/HBM binding fragment of Dkk, including Dkk-1, such as those depicted in Figure 6 or a mimetic of those fragments depicted in Figure 6. The Dkk modulating composition may also comprise a peptide aptamer of a Dkk interacting or Dkk-1 interacting protein, or a mimetic of such an aptamer. The invention further provides such a method wherein the composition comprises an aptamer of Dkk or Dkk-1, such as those depicted in Figure 3 (SEQ ID NOs:171-188), or a mimetic of such an aptamer. A composition of the present invention may modulate activity either by enhancing or inhibiting the binding of Dkk, including Dkk-1, to LRP5, LRP6,

or HBM or the binding of Dkk, including Dkk-1, to a Dkk interacting protein, such as those shown in Figure 5. The invention further provides such a method wherein the composition comprises a Dkk or Dkk-1 antibody or an antibody to a Dkk interacting or Dkk-1 interacting protein. A composition of the present invention may comprise an LRP5 peptide aptamer, such as OST262 (SEQ ID NO:208), Figures 4 (SEQ ID NOs:189-192 (particularly peptide 13 (SEQ ID NO:191)) and 13 (including SEQ ID NOs:204-214), or a mimetic of such an aptamer. Preferred compositions of the present invention also comprise LRP5 antibodies. It is a further aspect of the invention that such lipid-modulated diseases include a cardiac condition, atherosclerosis, familial lipoprotein lipase deficiency, familial apoprotein CII deficiency, familial type 3 hyperlipoproteinemia, familial hypercholesterolemia, familial hypertriglyceridemia, multiple lipoprotein-type hyperlipidemia, elevated lipid levels due to dialysis and/or diabetes, and an elevated lipid level of unknown etiology.

Bone disorders contemplated for treatment and/or diagnosis by the methods and compositions disclosed herein include a bone development disorder, a bone fracture, age related loss of bone, a chondrodystrophy, a drug-induced bone disorder, high bone turnover, hypercalcemia, hyperostosis, osteogenesis imperfecta, osteomalacia, osteomyelitis, osteoporosis, Paget's disease, osteoarthritis, and rickets.

It is a further object of the invention to provide a method of screening for compounds or compositions which modulates the interaction of Dkk with LRP5, LRP6, HBM, or a Dkk-binding fragment of LRP5, LRP6, or HBM comprising:

(a) exposing Dkk or a LRP5/LRP6/HBM binding fragment thereof to a compound; and

(b) determining whether said compound binds to Dkk or the LRP5/LRP6/HBM binding fragment thereof.

In a preferred embodiment, the Dkk is Dkk-1. In a particularly preferred embodiment, the binding of Dkk-1 to LRP5/LRP6/HBM is decreased.

It is a further object of the invention to provide a method of screening compounds or compositions which modulate the interaction of DKK with LRP5, LRP6, HBM, or a DKK-finding fragment thereof comprising:

(a) exposing DKK or a LRP5/LRP6/HBM binding fragment thereof to a compound; and,

(b) determining whether said compound modulates the interaction of Dkk with LRP5, LRP6, or HBM, or the Dkk-binding fragment of LRP5/LRP6/HBM.

In a preferred embodiment, the Dkk is Dkk-1. In a particularly preferred embodiment, the interaction of Dkk-1 with LRP5/LRP6/HBM is decreased.

It is a further object of the invention to provide a method of screening for compounds or compositions which modulates the interaction of Dkk with LRP5, LRP6, HBM, or a Dkk-binding fragment of LRP5, LRP6, or HBM comprising:

(a) exposing Dkk or a LRP5/LRP6/HBM binding fragment thereof to a compound;

(b) determining whether said compound binds to Dkk or the LRP5/LRP6/HBM binding fragment thereof; and,

(c) further determining whether said compound modulates the interaction of Dkk with LRP5, LRP6, or HBM, or the Dkk-binding fragment of LRP5/LRP6/HBM.

In preferred embodiments of such methods, Dkk or a biologically active fragment thereof is attached to a solid substrate. In an alternative embodiment of the invention, LRP5/LRP6/HBM, or a biologically active fragment thereof (such as the ligand binding domain), is exposed to the compound. Another aspect of the invention provides for compounds and compositions identified by the disclosed methods. A preferred embodiment of the invention provides that the compound screened in an afore-mentioned method is one or more proteins which interact with Dkk, particularly Dkk-1, as depicted in Figure 5, or a LRP5/LRP6/HBM-binding fragment thereof. Another preferred embodiment provides that the compound comprises a Dkk or Dkk-1 peptide aptamer, such as those depicted in Figure 3 (SEQ

ID NOs:171-188), or a mimetic of such aptamers. The compound may also comprise a peptide aptamer of a Dkk interacting or Dkk-1 interacting protein, or a mimetic of such an aptamer. The method further provides that the compound comprises a Dkk or Dkk-1 antibody or an antibody to a Dkk-1 interacting protein. The invention further provides that the compound may comprise an LRP5 peptide aptamer, such as OST262 (SEQ ID NO:208), Figure 4 (SEQ ID NOs:189-192) (particularly peptide 13 (SEQ ID NO:191)) and Figure 13 (including SEQ ID NOs:204-214), or a mimetic of such an aptamer. Preferred compounds of the present invention also comprise LRP5 antibodies.

It is a further object of the invention to provide a method of screening for compounds or compositions which modulate the interaction of Dkk and a Dkk interacting protein comprising:

(a) exposing a Dkk interacting proteins or a Dkk-binding fragment thereof to a compound; and,

(b) determining whether said compound binds to a Dkk interacting proteins or the Dkk-binding fragment thereof.

In a preferred embodiment, the Dkk is Dkk-1.

It is a further object of the invention to provide a method of screening for compounds or compositions which modulate the interaction of Dkk and a Dkk interacting protein comprising:

(a) exposing Dkk interacting protein(s) or a Dkk-binding fragment thereof to a compounds; and,

(b) determining whether said compound modulates the interaction of Dkk and Dkk interacting proteins.

It is a further object of the invention to provide a method of screening for compounds or compositions which modulate the interaction of Dkk and a Dkk interacting protein comprising:

(a) exposing a Dkk interacting proteins or a Dkk-binding fragment thereof to a compound;

(b) determining whether said compound binds to a Dkk interacting proteins or the Dkk-binding fragment thereof; and,

(c) further determining whether said compound modulates the interaction of Dkk and Dkk interacting proteins.

In a preferred embodiment, Dkk is Dkk-1.

In preferred embodiments of such methods, the Dkk interacting proteins, particularly Dkk-1 interacting proteins, or a Dkk-binding fragment thereof are attached to a solid substrate. Another aspect of the invention provides for compounds and compositions identified by the disclosed methods. A preferred embodiment provides that the compound comprises a Dkk or Dkk-1 peptide aptamer, such as those depicted in Figure 3 (SEQ ID NOs:171-188), or a mimetic of such aptamers. The compound may also comprise a peptide aptamer of a Dkk interacting or Dkk-1 interacting protein, or a mimetic of such an aptamer. The compound may also comprise an antibody to a Dkk interacting or Dkk-1 interacting protein.

It is another object of the invention to provide for a composition for treating bone mass disorders comprising a LRP5/LRP6/HBM modulating compound and a pharmaceutically acceptable excipient and/or carrier therefor. Preferred LRP5 (or LRP6 or HBM) modulating compounds include Dkk or Dkk-1 or a LRP5/LRP6/HBM binding fragment thereof. Also contemplated are compounds which comprise monoclonal or polyclonal antibodies or immunologically active fragments thereof which bind to Dkk, including Dkk-1, and a pharmaceutically acceptable excipient and/or carrier. Another preferred embodiment provides that the modulating compound comprises one or more Dkk interacting or Dkk-1 interacting proteins, or a biologically active fragment thereof. Also contemplated are compounds which comprise monoclonal or polyclonal antibodies or immunologically active fragments thereof which bind to Dkk interacting or Dkk-1 interacting proteins, or a biologically active fragment thereof, and a pharmaceutically acceptable excipient and/or carrier.

Another preferred embodiment provides that the modulating compound comprises an antisense, siRNA, and shRNA molecule which recognizes and binds to a nucleic acid encoding a Dkk interacting or Dkk-1 interacting protein. Another preferred embodiment provides that the modulating compound comprises a Dkk or Dkk-1 peptide aptamer, a mimetic of a Dkk or Dkk-1 peptide aptamer, a peptide aptamer of a Dkk interacting or Dkk-1 interacting protein, or a mimetic of such an aptamer. Another embodiment provides that the compound comprises an LRP5 peptide aptamer, such as OST262 (SEQ ID NO:208), Figure 4 (SEQ ID NOs:189-192) (particularly peptide) and Figure 13 (including SEQ ID NOs:204-214), or a mimetic of such an aptamer. Preferred compounds of the present invention also comprise LRP5 antibodies.

It is a further object of the invention to provide a pharmaceutical composition for treating a Dkk-mediated disease or condition comprising a compound which modulates Dkk activity and a carrier therefor, including pharmaceutically acceptable excipients. Such compositions include those wherein the compound comprises an antisense, siRNA, and shRNA molecule or an antibody which binds to Dkk, including Dkk-1, and thereby prevents it from interacting with LRP5, LRP6, or HBM. Other such compositions include one or more of Dkk interacting or Dkk-1 interacting proteins, such as those depicted in Figure 5, or a Dkk-binding fragment thereof, or a monoclonal or polyclonal antibody, or immunologically active fragment thereof, which binds to a Dkk interacting or Dkk-1 interacting protein or Dkk-binding fragment thereof. Other contemplated compositions include antisense, siRNA, and shRNA molecules which recognize and bind to a nucleic acid encoding a Dkk interacting or Dkk-1 interacting protein. Further contemplated compositions include Dkk and Dkk-1 peptide aptamers, such as those depicted in Figure 3 (SEQ ID NOs;171-188), mimetics of such aptamers, a peptide aptamer of a Dkk interacting or Dkk-1 interacting protein, or a mimetic of such an aptamer. Other contemplated compositions comprise an LRP5 peptide aptamer, such as OST262 (SEQ ID NO:208), Figure 4 (SEQ ID NOs:189-192) (particularly peptide 13 (SEQ ID NO:191))

and Figure 13 (including SEQ ID NO:204-214), or a mimetic of such an aptamer. Other preferred compositions of the present invention comprise LRP5 antibodies.

A further object of the invention to provide for a method of modulating the expression of a nucleic acid encoding a Dkk interacting or Dkk-1 interacting protein in an organism, such as those shown in Figure 5, comprising the step of administering to the organism an effective amount of composition which modulates the expression of a nucleic acid encoding a Dkk-1 interacting protein. In a preferred embodiment, said composition comprises an antisense, siRNA, or shRNA molecule which recognizes and binds to a nucleic acid encoding a Dkk interacting or Dkk-1 interacting protein.

One aspect of the invention provides for a method of modulating at least one activity of Dkk or a Dkk-1 interacting protein comprising administering an effective amount of a composition which modulates at least one activity of Dkk or a Dkk-1 interacting protein. The invention provides for a composition comprising a Dkk interacting or Dkk-1 interacting protein, such as those shown in Figure 5, or a biologically active fragment thereof. Other agents contemplated for this method are antisense, siRNA, or shRNA molecules which recognize and bind to a nucleic acid encoding a Dkk interacting or Dkk-1 interacting protein. The method further provides that the composition comprises a Dkk or Dkk-1 antibody or an antibody to a Dkk interacing or Dkk-1 interacting protein. In another preferred embodiment, the composition comprises a Dkk or Dkk-1 peptide aptamer, a mimetic of a Dkk or Dkk-1 peptide aptamer, a peptide aptamer of a Dkk interacting or Dkk-1 interacting protein, or a mimetic of such an aptamer. The method provides that a composition of the present invention may comprise an LRP5 peptide aptamer, such as OST262 (SEQ ID NO:208), Figure 4 (SEQ ID NO:189-192) (particularly peptide including (SEQ ID NO:191)) and Figure including (SEQ ID NOs:204-214), or a mimetic of such an aptamer. Preferred compositions of the present invention also comprise LRP5 antibodies. In a further preferred embodiment, the modulated Dkk activity is lipid modulation or bone mass modulation.

In all of the testing/screening embodiments of the present invention discussed below to obtain compounds or compositions which ultimately impact LRP5/LRP6/HBM signaling, one skilled in the art will recognize that HBM can be used as a control in the absence of a test sample or compound. Further, the effect of a test sample of compound on Wnt signaling through the interaction of Dkk with LRP5/LRP6/HBM does not necessarily require a direct measurement of an association or interaction of Dkk and LRP5/LRP6/HBM. Other positive phenotypes/activities established by the High Bone Mass phenotype or by using HBM as a control.

One aspect of the invention provides for a method of identifying binding partners for a Dkk protein comprising the steps of:

(a) exposing the Dkk protein(s) or a LRP5/LRP6 binding fragment thereof to a potential binding partner; and

(b) determining if the potential binding partner binds to a Dkk protein or the LRP5/LRP6 binding fragment thereof.

In a preferred embodiment, the Dkk is Dkk-1.

Another aspect of the invention is to provide for a method of identifying a compound that effects Dkk-mediated activity comprising

(a) providing a group of transgenic animals having (1) a regulatable one or more Dkk interacting protein genes, (2) a knock-out of one or more Dkk interacting protein genes, or (3) a knock-in of one or more Dkk interacting protein genes;

(b) providing a second group of control animals respectively for the group of transgenic animals in step (a); and

(c) exposing the transgenic animal group and the control animal group to a potential Dkk-modulating compound which modulates bone mass or lipid levels; and

(d) comparing the transgenic animal group and the control animal group and determining the effect of the compound on bone mass or lipid levels in the transgenic animals as compared to the control animals.

In a preferred embodiment, the Dkk is Dkk-1.

It is another aspect of the invention to provide for a method for determining whether a compound modulates a Dkk interacting protein, said method comprising the steps of:

(a) mixing the Dkk interacting protein or a Dkk-binding fragment thereof with the ligand binding domain of Dkk in the presence of said at least one compound;

(b) measuring the amount of said binding domain of Dkk bound to said Dkk interacting protein or the Dkk-binding fragment thereof as compared to a control without said at least one compound; and

(c) determining whether the compound reduces the amount of said binding domain of Dkk binding to said Dkk interacting protein or Dkk-binding fragment thereof.

In a preferred embodiment, the Dkk is Dkk-1.

In a preferred embodiment, the binding domain is attached to a solid substrate. The invention further provides for compounds identified by this method. In a preferred embodiment, the invention provides that the Dkk interacting or Dkk-1 interacting protein is detected by antibodies. In another preferred embodiment, the solid substrate is a microarray. Another preferred embodiment provides that the ligand binding domain of Dkk and/or Dkk interacting protein is fused or conjugated to a peptide or protein. The invention also provides that the compounds include Dkk

and Dkk-1 peptide aptamers, mimetics of Dkk and Dkk-1 peptide aptamers, Dkk and Dkk-1 interacting proteins peptide aptamers, or mimetics of such aptamers.

An aspect of the invention provides a composition comprising one or more polypeptide sequences of one or more Dkk-1 interacting proteins, or a biologically active fragment thereof, one or more Dkk proteins, or a biologically active fragment thereof, or LRP5/LRP6/HBM polypeptide sequences or a biologically active fragment thereof (for example, the ligand binding domain) and a pharmaceutically acceptable excipient and/or carrier. Another aspect of the invention provides that the composition comprises a Dkk or Dkk-1 antibody or an antibody to a Dkk interacting or Dkk-1 interacting protein and a pharmaceutically acceptable excipient. A composition of the present invention may comprise an LRP5 peptide aptamer, such as OST262 (SEQ ID NO:208), Figure 4 (SEQ ID NOs:189-192) (particularly peptide 13 (SEQ ID NO:191)) and Figure 13 (including SEQ ID NOs:204-214), or a mimetic of such an aptamer. A composition of the present invention may comprise a Dkk peptide aptamer, for example as shown in Figure 3 (SEQ ID NOs:171-188). Preferred compositions of the present invention also comprise LRP5 antibodies.

Another aspect of the invention is to provide an antibody or immunologically active antibody fragment which recognizes and binds to a Dkk-1 amino acid sequence selected from the group consisting of: Asn34-His266 (SEQ ID NO:110), Asn34-Cys245 (SEQ ID NO:111), Asn34-Lys182 (SEQ ID NO:112), Cys97-His266 (SEQ ID NO:113), Val139-His266 (SEQ ID NO:114), Gly183-His266 (SEQ ID NO:115), Cys97-Cys245 (SEQ ID NO:116), or Val139-Cys245 (SEQ ID NO:117) of human Dkk-1. Additional antibodies may bind to any of the sequences depicted in Figures 3 (SEQ ID NOs:171-188) and Figure 4 (SEQ ID NOs:189-192). Another aspect of the invention is to provide for polyclonal antibodies to one or more amino acid sequences: Peptide 1 -GNKYQTIDNYQPYPC (SEQ ID NO:118), Peptide 2 - LDGYSRRTTLSSKMYHTKGQEG (SEQ ID NO:119), Peptide 3 - RIQKDHHQASNSSRLHTCQRH (SEQ ID NO:120), Peptide 4 - RGEIEETITESFGND (SEQ ID NO:121), and Peptide 5 - EIFQRCYCGEGLSCRIQKD (SEQ ID NO: 122).

It is a further object of the invention to provide a nucleic acid encoding a Dkk protein, e.g. Dkk-1, a Dkk interacting or Dkk-1 interacting protein aptamer, or an LRP5 aptamer comprising a nucleic acid encoding a scaffold protein in-frame with the activation domain of Gal4 or LexA that is in-frame with a nucleic acid which encodes for a Dkk or Dkk-1 or Dkk interacting or Dkk-1 interacting protein amino acid sequence. Preferably the scaffold protein is thioredoxin (trxA), S1 nuclease from *Staphylococcus* or M13. Other preferable embodiments include Dkk-1 amino acid sequences selected from Figure 6.

It is yet a further object of the invention to provide a composition comprising a polypeptide sequence of Figure 3 (SEQ ID NOs:171-188), Figure 4 (SEQ ID NO:189-192), or of Dkk-1 interacting proteins identified in Figure 5 and a pharmaceutically acceptable excipient and/or carrier.

Another aspect of the invention includes a method of detecting the modulatory activity of a compound on the binding interaction of a first peptide and a second peptide of a peptide binding pair that bind through extracellular interaction in their natural environment, comprising:

(i) culturing at least one eukaryotic cell, wherein the eukaryotic cell comprises;

a) a nucleotide sequence encoding a first heterologous fusion protein comprising the first peptide or a segment thereof joined to a DNA binding domain of a transcriptional activation protein;

b) a nucleotide sequence encoding a second heterologous fusion protein comprising the second peptide or a segment thereof joined to a transcriptional activation domain of a transcriptional activation protein;

wherein binding of the first peptide or segment thereof and the second peptide or segment thereof reconstitutes a transcriptional activation protein; and

c) a reporter element activated under positive transcriptional control of the reconstituted transcriptional activation protein, wherein expression of the reporter element produces a selected phenotype;

(ii) incubating a compound with the eukaryotic cell under conditions suitable to detect the selected phenotype; and

(iii) detecting the ability of the compound to affect the binding interaction of the peptide binding pair by determining whether the compound affects the expression of the reporter element which produces the selected phenotype;

wherein (1) said first peptide is a Dkk peptide and said second peptide is a peptide selected from LRP5, HBM, LRP6, and the Dkk-binding portion of LRP5/LRP6/HBM or (2) said first peptide is a Dkk-interacting protein or the Dkk-binding fragment thereof, and said second peptide is a Dkk peptide.

In one embodiment, the eukaryotic cell is a yeast cell. In a preferred embodiment, the yeast cell is *Saccharomyces*. In a particularly preferred embodiment, the *Saccharomyces* cell is *Saccharomyces cerevisiae*. The invention further provides that the compound may comprise a Dkk interacting or Dkk-1 interacting protein, or a biologically active fragment thereof. In one embodiment, the Dkk interacting or Dkk-1 interacting protein, or a Dkk-binding fragment thereof, is added directly to the assay. In another embodiment, the Dkk interacting or Dkk-1 interacting protein, or a Dkk-binding fragment thereof, is recombinantly expressed by the eukaryotic cell in addition to the first and second peptides. In a preferred embodiment the compound comprises a Dkk or Dkk-1 aptamer, a mimetic of a Dkk or Dkk-1 peptide aptamer, a Dkk interacting or Dkk-1 interacting protein aptamer, or a mimetic of a Dkk-1 interacting protein aptamer. Other preferred embodiments provide that the compound comprises an LRP5 peptide aptamer, such as OST262 (SEQ ID NO:208), Figure 4 (SEQ ID NOs:189-192) (particularly peptide 13 (SEQ ID NO:191) and Figure 13 (including SEQ ID NOs:204-214), or a mimetic of such an aptamer. Alternatively, the present invention also provides that the compound may

comprise LRP5 antibodies or Dkk antibodies. In another embodiment, the yeast cell further comprises at least one endogenous nucleotide sequence selected from the group consisting of a nucleotide sequence encoding the DNA binding domain of a transcriptional activation protein, a nucleotide sequence encoding the transcriptional activation domain of a transcriptional activation protein, and a nucleotide sequence encoding the reporter element, wherein at least one of the endogenous nucleotide sequences is inactivated by mutation or deletion. In another embodiment, the peptide binding pair comprises a ligand and a receptor to which the ligand binds. In one embodiment, the transcriptional activation protein is Gal4, Gcn4, Hap1, Adr1, Swi5, Ste12, Mcm1, Yap1, Ace1, Ppr1, Arg81, Lac9, Qa1F, VP16, or a mammalian nuclear receptor. In another embodiment, at least one of the heterologous fusion proteins is expressed from an autonomously-replicating plasmid. In one embodiment, the DNA binding domain comprises a heterologous DNA-binding domain of a transcriptional activation protein. In a preferred embodiment, the DNA binding protein is selected from the group consisting of a mammalian steroid receptor and bacterial LexA protein. In another embodiment, the reporter element is selected from the group consisting of *lacZ*, a polynucleotide encoding luciferase, a polynucleotide encoding green fluorescent protein (GFP), and a polynucleotide encoding chloramphenicol acetyltransferase. In a particularly preferred embodiment, the reporter element is lacZ

The invention further provides for a rescue screen for detecting the activity of a compound for modulating the binding interaction of a first peptide and a second peptide of a peptide binding pair, comprising:

(i) culturing at least one yeast cell, wherein the yeast cell comprises;

a) a nucleotide sequence encoding a first heterologous fusion protein comprising the first peptide or a segment thereof joined to a DNA binding domain of a transcriptional activation protein;

b) a nucleotide sequence encoding a second heterologous

fusion protein comprising the second peptide or a segment thereof joined to a transcriptional activation domain of a transcriptional activation protein;

wherein binding of the first peptide or segment thereof and the second peptide or segment thereof reconstitutes a transcriptional activation protein; and

c) a reporter element activated under positive transcriptional control of the reconstituted transcriptional activation protein, wherein expression of the reporter gene prevents exhibition of a selected phenotype;

(ii) incubating a compound with the yeast cell under conditions suitable to detect the selected phenotype; and

(iii) detecting the ability of the compound to affect the binding interaction of the peptide binding pair by determining whether the compound affects the expression of the reporter element which prevents exhibition of the selected phenotype,

wherein said first peptide is a Dkk peptide and said second peptide is a peptide selected from LRP5, HBM, LRP6 and a Dkk-binding fragment of LRP5/LRP6/HBM.

In a preferred embodiment, the invention provides that the yeast cell is *Saccharomyces*. In a particularly preferred embodiment, the *Saccharomyces* cell is *Saccharomyces cerevisiae*. In one embodiment, the compound comprises one or more Dkk interacting or Dkk-1 interacting proteins, or a Dkk-binding fragment thereof. Compounds used in the present invention may comprise an LRP5 peptide aptamer, such as OST262 (SEQ ID NO:208), Figure 4 (SEQ ID NOs:189-192) (particularly peptide 13 (SEQ ID NO:191)) and Figure 13 (including SEQ ID NOs:204-214), or a mimetic of such an aptamer. Alternatively, the compound may comprise LRP5 antibodies or Dkk antibodies. In another embodiment, the yeast cell further comprises at least one endogenous nucleotide sequence selected from the group consisting of a nucleotide sequence encoding the DNA binding domain of a

transcriptional activation protein, a nucleotide sequence encoding the transcriptional activation domain of a transcriptional activation protein, and a nucleotide sequence encoding the reporter gene, wherein at least one of the endogenous nucleotide sequences is inactivated by mutation or deletion. In another embodiment, the transcriptional activation protein is Gal4, Gcn4, Hap1, Adr1, Swi5, Ste12, Mcm1, Yap1, Ace1, Ppr1, Arg81, Lac9, Qa1F, VP16, or a mammalian nuclear receptor. In one embodiment, at least one of the heterologous fusion proteins is expressed from an autonomously-replicating plasmid. In another embodiment, the DNA binding domain is a heterologous DNA-binding domain of a transcriptional activation protein.

The invention also provides for a rescue screen for detecting the modulatory activity of a compound on the binding interaction of a first peptide and a second peptide of a peptide binding pair, comprising:

(i) culturing at least one yeast cell, wherein the yeast cell comprises;

a) a nucleotide sequence encoding a first heterologous fusion protein comprising the first peptide or a segment thereof joined to a DNA binding domain of a transcriptional activation protein;

b) a nucleotide sequence encoding a second heterologous fusion protein comprising the second peptide or a segment thereof joined to a transcriptional activation domain of a transcriptional activation protein;

wherein binding of the first peptide or segment thereof and the second peptide or segment thereof reconstitutes a transcriptional activation protein; and

c) a reporter element activated under positive transcriptional control of the reconstituted transcriptional activation protein, wherein expression of the reporter element prevents exhibition of a selected phenotype;

(ii) incubating a compound with the yeast cell under conditions suitable to detect the selected phenotype; and

(iii) detecting the ability of the compound to affect the binding interaction of the peptide binding pair by determining whether the compound affects the expression of the reporter element which prevents exhibition of the selected phenotype,

wherein said first peptide is a Dkk interacting or Dkk-1 interacting protein peptide and said second peptide is a Dkk or Dkk-1 peptide.

In a preferred embodiment of the rescue screen, the yeast cell is *Saccharomyces*. In a particularly preferred embodiment, the *Saccharomyces* cell is *Saccharomyces cerevisiae*. In another embodiment, the yeast cell further comprises at least one endogenous nucleotide sequence selected from the group consisting of a nucleotide sequence encoding the DNA binding domain of a transcriptional activation protein, a nucleotide sequence encoding the transcriptional activation domain of a transcriptional activation protein, and a nucleotide sequence encoding the reporter gene, wherein at least one of the endogenous nucleotide sequences is inactivated by mutation or deletion. In one embodiment, the transcriptional activation protein is Gal4, Gcn4, Hap1, Adr1, Swi5, Ste12, Mcm1, Yap1, Ace1, Ppr1, Arg81, Lac9, Qa1F, VP16, or a mammalian nuclear receptor. In another embodiment of the rescue screen, at least one of the heterologous fusion proteins is expressed from an autonomously-replicating plasmid. In another embodiment, the DNA binding domain is a heterologous DNA-binding domain of a transcriptional activation protein.

The invention also provides for a method for identifying potential compounds which modulate Dkk activity comprising:

a) measuring the effect on binding of one or more Dkk interacting protein, or a Dkk-binding fragment thereof, with Dkk or a LRP5/LRP6/HBM binding fragment thereof in the presence and absence of a compound; and

b) identifying as a potential Dkk modulatory compound a compound which modulates the binding between one or more Dkk interacting proteins or Dkk-binding fragment thereof and Dkk or LRP5/LRP6/HBM fragment thereof.

In a preferred embodiment, the Dkk is Dkk-1.

The invention further provides for any of the Dkk peptide aptamers of Figure 3 (SEQ ID NOs:171-188). The invention also provides for any of the LRP peptide aptamers of Figure 4 (SEQ ID NOs:189-192).

Another aspect of the invention provides for a method of identifying agents which modulate the interaction of Dkk with the Wnt signaling pathway comprising:

(a) injecting mRNA encoding Dkk and an agent into a *Xenopus* blastomere;

(b) assessing axis duplication or analyzing marker gene expression; and

(c) identifying agents which elicit changes in axis duplication or marker gene expression as agents which modulate the interaction of Dkk with the Wnt signaling pathway. Wherein the agent may be chosen from among mRNA encoding Dkk interacting proteins, fragments thereof, siRNA, shRNA, antisense nucleotides, and antibodies. In a preferred embodiment, Dkk is Dkk-1. In a further embodiment, mRNA of HBM, LRP5/6, any Wnt (including Wnt1-Wnt19, particularly Wnt1, Wnt3, Wnt3a, and Wnt10b), Wnt antagonist, or combination of these is co-injected into the *Xenopus* blastomere. In another embodiment, the marker gene analyzed could include Siamois, Xnr3, slug, Xbra, HNK-1, endodermin, Xlhbox8, BMP2, BMP4, XLRP6, EF-1, or ODC.

The present invention provides for a method for identifying agents which modulate the interaction of Dkk with the Wnt signaling pathway comprising:

(a) transfecting cells with constructs encoding Dkk and potential Dkk interacting proteins, mRNA fragments thereof, siRNA, shRNA, or antisense, antibodies to LRP5/HBM/LRP6/Dkk/Dkk-interacting protein;

(b) assessing changes in expression of a reporter gene linked to a Wnt-responsive promoter; and,

(c) identifying as a Dkk interacting protein any protein which alters reporter gene expression compared with cells transfected with a Dkk construct alone. In a further preferred embodiment, the cells may be HOB-03-CE6, HEK293, or U2OS cells.

In alternative embodiments, the Wnt-responsive promoter is TCF or LEF. In other preferred embodiments, the cells are co-transfected with CMV beta-galactosidase or tk-Renilla.

The present invention further provides for a LRP5/HBM monoclonal or polyclonal antibody to one or more peptides of amino acid sequences MYWTDWVETPRIE (SEQ ID NO:123), MYWTDWGETPRIE (SEQ ID NO:124), KRTGGKRKEILSA (SEQ ID NO:125), ERVEKTTGDKRTRIQGR (SEQ ID NO:126), or KQQCDSFPDCIDGSDE (SEQ ID NO:127).

Additionally, the present invention provides a method for identifying compounds which modulate Dkk and LRP5/LRP6/HBM interactions comprising:

(a) immobilizing LRP5/LRP6/HBM to a solid surface; and

(b) treating the solid surface with a secreted Dkk protein or a secreted epitope-tagged Dkk and a test compound; and

(c) determining whether the compound regulates binding between Dkk and LRP5/LRP6/HMB using antibodies to Dkk or the epitope tag or by directly measuring activity of an epitope tag.

In one embodiment, the Dkk is Dkk-1. In a preferred embodiment, the epitope tag is alkaline phosphatase, histidine, myc, or a V5 tag.

Another embodiment of the present invention provides for a method for identifying compounds which modulate Dkk and LRP5/LRP6/HBM interactions comprising:

(a) creating an LRP5, LRP6, or HBM fluorescent fusion protein using a first fluorescent tag;

(b) creating a Dkk fusion protein comprising a second fluorescent tag;

(c) adding a test compound; and,

(d) assessing changes in the ratio of fluorescent tag emissions using Fluorescence Resonance Energy Transfer (FRET) or Bioluminescent Resonance Energy Transfer (BRET) to determine whether the compound modulates Dkk and LRP5/LRP6/HBM interactions.

In a preferred embodiment, the Dkk is Dkk-1.

The present invention also provides for a method of diagnosing low or high bone mass and/or low or high lipid levels in a subject comprising examining expression of Dkk, LRP5, LRP6, HBM or HBM-like variant in the subject and determining whether Dkk, LRP5, LRP6, or HBM or a HBM-like variant is over- or under-expressed to determine whether subject has (a) high or low bone mass and/or (b) high or low lipid levels.

The invention further provides for a transgenic animal wherein Dkk is knocked out in a tissue-specific fashion. In a preferred embodiment, the Dkk is Dkk-1. In one preferred embodiment, the tissue specificity is bone tissue. In another preferred embodiment, the tissue specificity is liver or other tissues or cells involved in regulating lipid metabolism or cancer tissue.

The present invention further provides a method of screening for compounds which modulate the interaction of Dkk with LRP5, LRP6, or HBM comprising:

(a) exposing LRP5, LRP6, or HBM, or a Dkk-binding fragment of LRP5, LRP6, or HBM to a compound; and
(b) determining whether said compound bound to LRP5, LRP6, or HBM or the Dkk-binding fragment of LRP5, LRP6, or HBM and further determining whether said compound modulates the interaction of Dkk and LRP5, LRP6, or HBM.

In one embodiment, the Dkk is Dkk-1. In a preferred embodiment, the compound comprises an LRP5 peptide aptamer. Other preferred compositions include the peptide aptamer, OST262 (SEQ ID NO:208), Figure 4 (SEQ ID NOs:189-192) (particularly peptide 13 (SEQ ID NO:191) and Figure 13 (including SEQ ID NOs:204-214), or a mimetic of such an aptamer, and an LRP5 antibody.

The present invention also provides a method for identifying compounds which modulate Dkk and LRP5/LRP6/HBM interactions comprising:

(a) immobilizing LRP5/LRP6/HBM to a solid surface; and
(b) treating the solid surface with a secreted Dkk protein or a secreted epitope-tagged Dkk and a test compound; and

(c) determining whether the compound regulates binding between Dkk and LRP5/LRP6/HBM using antibodies to Dkk or the epitope tag or by directly measuring activity of an epitope tag. In a preferred embodiment, the epitope tag is alkaline phosphatase, histidine, myc or a V5 tag.

In a preferred embodiment, the Dkk is Dkk-1.

The invention also provides for a method for identifying compounds which modulate the interaction of Dkk with the Wnt signaling pathway comprising:

(a) transfecting cells with constructs containing Dkk and Wnt proteins;

(b) assessing changes in expression of a reporter element linked to a Wnt-responsive promoter; and

(c) identifying as a Dkk/Wnt interaction modulating compound any compound which alters reporter gene expression compared with cells transfected with a Dkk construct alone.

In one embodiment, the Dkk is Dkk-1. In another embodiment, the Wnt is any of Wnt1-Wnt19. In a preferred embodiment, the Wnt is Wnt1, Wnt3, Wnt3a, or Wnt10b. In a particularly preferred embodiment, the Wnt construct contains Wnt3a. In another particularly preferred embodiment, the Wnt construct contains Wnt1. In another preferred embodiment, the Wnt construct encodes for a Wnt that signals through the canonical Wnt pathway. In a particularly preferred embodiment, both Wnt3a and Wnt1 constructs are co-transfected into the cells. In another embodiment, the cells may be U2-OS, HOB-03-CE6, or HEK293 cells. In another embodiment, the reporter element used is TCF-luciferase, tk-Renilla, or a combination thereof.

The invention also provides for a method of testing compounds that modulate Dkk-mediated activity in a mammal comprising:

(a) providing a group of transgenic animals having (1) a regulatable one or more Dkk genes, (2) a knock-out of Dkk genes, or (3) a knock-in of one or more Dkk genes;

(b) providing a second group of control animals respectively for the group of transgenic animals in step (a); and

(c) exposing the transgenic animal group and control animal group to a potential Dkk-modulating compound which modulates bone mass or lipid levels; and

(d) comparing the transgenic animals and the control group of animals and determining the effect of the compound on bone mass or lipid levels in the transgenic animals as compared to the control animals.

In a preferred embodiment, the Dkk is Dkk-1.

The invention further provides variants of LRP5 which demonstrate HBM biological activity, i.e., that are "HBM-like." In preferred embodiments, variants G171F, M282V, G171K, G171Q, A65V, G171V, G171I, and A214V of LRP5 are provided. The invention further provides for the use any of these variants in the forgoing methods.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows a schematic of the components of the Wnt signal transduction pathway. Schematic obtained from: http://www.stanford.edu/~rnusse/pathways/cell2.html

Figure 2 (A-C) show bait sequences (SEQ ID NOs:168-170) utilized in yeast two hybrid (Y2H) screens for protein-protein interactions.

Figure 3 shows a table of peptide aptamer insert sequences (SEQ ID NOs: 171-192) identified in Y2H screen with a Dkk-1 bait sequence.

Figure 4 shows a table of peptide aptamer insert sequences identified in a Y2H screen using a LRP5 ligand binding domain bait sequence.

Figure 5 shows a table of proteins identified in a Y2H screen using a Dkk-1 bait sequence. These proteins are identified by both their nucleic acid and amino acid accession numbers.

Figure 6 shows the results of a minimum interaction domain mapping screen of Dkk-1 with LRP5. At the top, a map of Dkk-1 showing the location of the signal

sequence, and cysteine rich domains 1 and 2. Below, the extent of domains examined using LRP5 LBD baits, LBD1 and LBD4, of Figure 2. To the right, scoring of the binding results observed in the experiment.

Figure 7 shows a diagram of the *Xenopus* Embryo Assay for Wnt activity.

Figure 8 shows the effects of Zmax/LRP5 and HBM on Wnt signaling in the *Xenopus* embryo assay.

Figure 9 shows the effects of Zmax/LRP5 and HBM on induction of secondary axis formation in the *Xenopus* embryo assay.

Figure 10 shows the effects of human Dkk-1 on the repression of the canonical Wnt pathway.

Figure 11 shows the effects of human Dkk-1 on Zmax/LRP5 and HBM-mediated Wnt signaling.

Figure 12 shows pcDNA3.1 construct names with nucleotide sequences (including SEQ ID NOs:193-203) for LRP5-binding peptide aptamers, Dkk-1 peptides and control constructs.

Figure 13 shows the amino acid sequences (including SEQ ID NOs:204-214) for the corresponding LRP5-binding peptides, Dkk-1 peptide aptamers and control constructs in Figure 12.

Figure 14 shows the effects of Dkk-1 and Dkk-2 on Wnt1 signaling with coreceptors LRP5, HBM, and LRP6 in HOB03CE6 cells.

Figure 15 shows the effects of Dkk-1 and Dkk-2 on Wnt3a signaling with coreceptors LRP5, HBM, and LRP6 in HOB03CE6 cells.

Figure 16 demonstrates that the LRP5-LBD peptide aptamer 262 activates Wnt signaling in the presence of Wnt3a in U2OS cells.

Figure 17 shows the differential binding of an antibody generated to a sequence (a.a. 165-177) containing the HBM mutation in LRP5 in LRP5 and HBM virus-infected cells.

Figure 18 shows data generated from a Y2H interaction trap where a mutant Dkk-1 (C220A) is unable to bind to LRP5 and demonstrating the window of capability of detecting small molecule effects on LRP and Dkk interactions.

Figure 19 shows that Dkk-1 represses Wnt3a-mediated Wnt signaling in U2OS bone cells using the cell-based reporter gene assay for high throughput screening.

Figure 20 demonstrates that Wnt1-HBM generated signaling is not efficiently inhibited by Dkk-1 in U2OS bone cells while LRP5 and LRP6-mediated signaling are using the cell-based reporter gene assay for high throughput screening.

Figure 21 shows that the TCF signal in the cell-based reporter gene assay for high throughput screening can be modulated by Dkk-1 and Dkk-1-AP without Wnt DNA transfection.

Figure 22 shows the morphological results in the Xenopus assay using aptamers 261 and 262 from the LRP5-LBD to activate Wnt signaling.

Figure 23 demonstrates that LRP5-LBD aptamers 261 and 262 induce Wnt signaling over other LRP5 aptamers.

Figure 24 shows that the mutation G171F in LRP5 produces a greater activation of the Wnt pathway than LRP5 which is consistent with HBM activity.

Figure 25 shows that the mutation M282V in LRP5 produces an activation of the Wnt pathway which is consistent with HBM activity in U2OS cells.

Figure 26 shows the amino acid sequence of the various peptides of dkk-1 selected to generate polyclonal antibodies, their relationship to the Dkk-1 amino acid sequence and identities of polyclonal antibodies generated.

Figure 27 shows a Western blot demonstrating that polyclonal antibody #5521 to amino acids 165-186 of Dkk-1 was able to detect Dkk1-V5 and Dkk1-AP from conditioned medium.

Figure 28 shows a Western blot demonstrating that polyclonal antibody #74397 to amino acids 147-161 was able to detect Dkk1-V5 in both conditioned medium and immunoprecipitated conditioned medium.

DETAILED DESCRIPTION OF THE INVENTION

1. Definitions

In general, terms in the present application are used consistent with the manner in which those terms are understood in the art. To aid in the understanding of the specification and claims, the following definitions are provided.

"Gene" refers to a DNA sequence that encodes through its template or messenger RNA a sequence of amino acids characteristic of a specific peptide. The term "gene" includes intervening, non-coding regions, as well as regulatory regions, and can include 5' and 3' ends.

By "nucleic acid" is meant to include single stranded and double stranded nucleic acids including, but not limited to DNAs, RNAs (e.g., mRNA, tRNAs, siRNAs), cDNAs, recombinant DNA (rDNA), rRNAs, antisense nucleic acids, oligonucleotides, and oligomers, and polynucleotides. The term may also include hybrids such as triple stranded regions of RNA and/or DNA or double stranded RNA:DNA hybrids. The term also is contemplated to include modified nucleic acids such as, but not limited to biotinylated nucleic acids, tritylated nucleic acids, fluorophor labeled nucleic acids, inosine, and the like.

"Gene sequence" refers to a nucleic acid molecule, including DNA which contains a non-transcribed or non-translated sequence, which comprises a gene. The term is also intended to include any combination of gene(s), gene fragment(s), non-transcribed sequence(s) or non-translated sequence(s) which are present on the same DNA molecule.

The nucleic acid sequences of the present invention may be derived from a variety of sources including DNA, cDNA, synthetic DNA, synthetic RNA or combinations thereof. Such sequences may comprise genomic DNA which may or may not include naturally occurring introns. Moreover, such genomic DNA may be obtained in association with promoter regions and/or poly (A) sequences. The sequences, genomic DNA or cDNA may be obtained in any of several ways. Genomic DNA can be extracted and purified from suitable cells by means well

known in the art. Alternatively, mRNA can be isolated from a cell and used to produce cDNA by reverse transcription or other means.

"cDNA" refers to complementary or copy DNA produced from an RNA template by the action of RNA-dependent DNA polymerase (reverse transcriptase). Thus, a "cDNA clone" means a duplex DNA sequence for which one strand is complementary to an RNA molecule of interest, carried in a cloning vector or PCR amplified. cDNA can also be single stranded after first strand synthesis by reverse transcriptase. In this form, it is a useful PCR template and does not need to be carried in a cloning vector. This term includes genes from which the intervening sequences have been removed. Thus, the term "gene", as sometimes used generically, can also include nucleic acid molecules comprising cDNA and cDNA clones.

"Recombinant DNA" means a molecule that has been engineered by splicing *in vitro* a cDNA or genomic DNA sequence or altering a sequence by methods such as PCR mutagenesis.

"Cloning" refers to the use of *in vitro* recombination techniques to insert a particular gene or other DNA sequence into a vector molecule. In order to successfully clone a desired gene, it is necessary to use methods for generating DNA fragments, for joining the fragments to vector molecules, for introducing the composite DNA molecule into a host cell in which it can replicate, and for selecting the clone having the target gene from amongst the recipient host cells.

"cDNA library" refers to a collection of recombinant DNA molecules containing cDNA inserts which together comprise the entire or a partial repertoire of genes expressed in a particular tissue or cell source. Such a cDNA library can be prepared by methods known to one skilled in the art and described by, for example, Cowell and Austin, "cDNA Library Protocols," *Methods in Molecular Biology* (1997).

"Cloning vehicle" refers to a plasmid or phage DNA or other DNA sequence which is able to replicate in a host cell. This term can also include artificial chromosomes such as BACs and YACs. The cloning vehicle is characterized by one or more endonuclease recognition sites at which such DNA sequences may be cut in

a determinable fashion without loss of an essential biological function of the DNA, which may contain a marker suitable for use in the identification of transformed cells.

"Expression" refers to the process comprising transcription of a gene sequence and subsequent processing steps, such as translation of a resultant mRNA to produce the final end product of a gene. The end product may be a protein (such as an enzyme or receptor) or a nucleic acid (such as a tRNA, antisense RNA, or other regulatory factor). The term "expression control sequence" refers to a sequence of nucleotides that control or regulate expression of structural genes when operably linked to those genes. These include, for example, the lac systems, the trp system, major operator and promoter regions of the phage lambda, the control region of fd coat protein and other sequences known to control the expression of genes in prokaryotic or eukaryotic cells. Expression control sequences will vary depending on whether the vector is designed to express the operably linked gene in a prokaryotic or eukaryotic host, and may contain transcriptional elements such as enhancer elements, termination sequences, tissue-specificity elements and/or translational initiation and termination sites.

"Expression vehicle" refers to a vehicle or vector similar to a cloning vehicle but which is capable of expressing a gene which has been cloned into it, after transformation into a host. The cloned gene is usually placed under the control of (i.e., operably linked to) an expression control sequence.

"Operator" refers to a DNA sequence capable of interacting with the specific repressor, thereby controlling the transcription of adjacent gene(s).

"Promoter" refers to a DNA sequence that can be recognized by an RNA polymerase. The presence of such a sequence permits the RNA polymerase to bind and initiate transcription of operably linked gene sequences.

"Promoter region" is intended to include the promoter as well as other gene sequences which may be necessary for the initiation of transcription. The presence of a promoter region is sufficient to cause the expression of an operably linked gene sequence. The term "promoter" is sometimes used in the art to generically indicate a promoter region. Many different promoters are known in the art which direct

expression of a gene in a certain cell types. Tissue-specific promoters can comprise nucleic acid sequences which cause a greater (or decreased) level of expression in cells of a certain tissue type.

"Operably linked" means that the promoter controls the initiation of expression of the gene. A promoter is operably linked to a sequence of proximal DNA if upon introduction into a host cell the promoter determines the transcription of the proximal DNA sequence(s) into one or more species of RNA. A promoter is operably linked to a DNA sequence if the promoter is capable of initiating transcription of that DNA sequence.

"Prokaryote" refers to all organisms without a true nucleus, including bacteria.

"Eukaryote" refers to organisms and cells that have a true nucleus, including mammalian cells.

"Host" includes prokaryotes and eukaryotes, such as yeast and filamentous fungi, as well as plant and animal cells. The term includes an organism or cell that is the recipient of a replicable expression vehicle.

The term "animal" is used herein to include all vertebrate animals, except humans. It also includes an individual animal in all stages of development, including embryonic and fetal stages. Preferred animals include higher eukaryotes such as avians, rodents (e.g., mice, rabbits, rats, chinchillas, guinea pigs, hamsters and the like), and mammals. Preferred mammals include bovine, equine, feline, canine, ovine, caprine, porcine, buffalo, humans, and primates.

A "transgenic animal" is an animal containing one or more cells bearing genetic information received, directly or indirectly, by deliberate genetic manipulation or by inheritance from a manipulated progenitor at a subcellular level, such as by microinjection or infection with a recombinant viral vector (e.g., adenovirus, retrovirus, herpes virus, adeno-associated virus, lentivirus). This introduced DNA molecule may be integrated within a chromosome, or it may be extra-chromosomally replicating DNA.

"Embryonic stem cells" or "ES cells" as used herein are cells or cell lines usually derived from embryos which are pluripotent meaning that they are un-

differentiated cells. These cells are also capable of incorporating exogenous DNA by homologous recombination and subsequently developing into any tissue in the body when incorporated into a host embryo. It is possible to isolate pluripotent cells from sources other than embryonic tissue by methods which are well understood in the art.

Embryonic stem cells in mice have enabled researchers to select for transgenic cells and perform gene targeting. This allows more genetic engineering than is possible with other transgenic techniques. For example, mouse ES cells are relatively easy to grow as colonies *in vitro*. The cells can be transfected by standard procedures and transgenic cells clonally selected by antibiotic resistance. See, for example, Doetschman *et al.*., 1994, *Gene transfer in embryonic stem cells.* In Pinkert (Ed.) Transgenic Animal Technology: A Laboratory Handbook. Academic Press Inc., New York, pp.115-146. Furthermore, the efficiency of this process is such that sufficient transgenic colonies (hundreds to thousands) can be produced to allow a second selection for homologous recombinants. Mouse ES cells can then be combined with a normal host embryo and, because they retain their potency, can develop into all the tissues in the resulting chimeric animal, including the germ cells. The transgenic modification can then be transmitted to subsequent generations.

Methods for deriving embryonic stem (ES) cell lines *in vitro* from early preimplantation mouse embryos are well known. See for example, Evans *et al.*, 1981 *Nature* 29: 154-6 and Martin, 1981, *Proc. Nat. Acad. Sci. USA*, 78: 7634-8. ES cells can be passaged in an undifferentiated state, provided that a feeder layer of fibroblast cells or a differentiation inhibiting source is present.

The term "somatic cell" indicates any animal or human cell which is not a sperm or egg cell or is capable of becoming a sperm or egg cell. The term "germ cell" or "germ-line cell" refers to any cell which is either a sperm or egg cell or is capable of developing into a sperm or egg cell and can therefore pass its genetic information to offspring. The term "germ cell-line transgenic animal" refers to a transgenic animal in which the genetic information was incorporated in a germ line cell, thereby conferring the ability to transfer the information to offspring. If such

offspring in fact possess some or all of that information, then they, too, are transgenic animals.

The genetic alteration of genetic information may be foreign to the species of animal to which the recipient belongs, or foreign only to the particular individual recipient. In the last case, the altered or introduced gene may be expressed differently than the native gene.

"Fragment" of a gene refers to any portion of a gene sequence. A "biologically active fragment" refers to any portion of the gene that retains at least one biological activity of that gene. For example, the fragment can perhaps hybridize to its cognate sequence or is capable of being translated into a polypeptide fragment encoded by the gene from which it is derived.

"Variant" refers to a gene that is substantially similar in structure and biological activity or immunological characteristics to either the entire gene or to a fragment of the gene. Provided that the two genes possess a similar activity, they are considered variant as that term is used herein even if the sequence of encoded amino acid residues is not identical. Preferentially, as used herein (unless otherwise defined) the variant is one of LRP5, HBM or LRP6. The variant preferably is one that yields an HBM-like phenotype (i.e., enhances bones mass and/or modulates lipid levels). These variants include missense mutations, single nucleotide polymorphisms (SNPs), mutations which result in changes in the amino acid sequence of the protein encoded by the gene or nucleic acid, and combinations thereof, as well as com in the exon domains of the *HBM* gene and mutations in LRP5 or LRP6 which result in an HBM like phenotype.

"Amplification of nucleic acids" refers to methods such as polymerase chain reaction (PCR), ligation amplification (or ligase chain reaction, LCR) and amplification methods based on the use of Q-beta replicase. These methods are well known in the art and described, for example, in U.S. Patent Nos. 4,683,195 and 4,683,202. Reagents and hardware for conducting PCR are commercially available. Primers useful for amplifying sequences from the HBM region are preferably complementary to, and hybridize specifically to sequences in the HBM region or in

regions that flank a target region therein. HBM sequences generated by amplification may be sequenced directly. Alternatively, the amplified sequence(s) may be cloned prior to sequence analysis.

"Antibodies" may refer to polyclonal and/or monoclonal antibodies and fragments thereof, and immunologic binding equivalents thereof, that can bind to the HBM proteins and fragments thereof or to nucleic acid sequences from the HBM region, particularly from the HBM locus or a portion thereof. Preferred antibodies also include those capable of binding to LRP5, LRP6 and HBM variants. The term antibody is used both to refer to a homogeneous molecular entity, or a mixture such as a serum product made up of a plurality of different molecular entities. Proteins may be prepared synthetically in a protein synthesizer and coupled to a carrier molecule and injected over several months into rabbits. Rabbit sera is tested for immunoreactivity to the HBM protein or fragment. Monoclonal antibodies may be made by injecting mice with the proteins, or fragments thereof. Monoclonal antibodies will be screened by ELISA and tested for specific immunoreactivity with HBM protein or fragments thereof. Harlow *et al.*, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1988) and *Using Antibodies: A Laboratory Manual*, Harlow, Ed and Lane, David (Cold Spring Harbor Press, 1999). These antibodies will be useful in assays as well as pharmaceuticals. By "antibody" is meant to include but not limited to polyclonal, monoclonal, chimeric, human, humanized, bispecific, multispecific, primatized™ antibodies.

"HBM protein" refers to a protein that is identical to a Zmax1 (LRP5) protein except that it contains an alteration of glycine 171 to a valine. An HBM protein is defined for any organism that encodes a Zmax1 (LRP5) true homolog. For example, a mouse HBM protein refers to the mouse Zmax1 (LRP5) protein having the glycine 170 to valine substitution.

By "HBM-like" is meant a variant of LRP5, LRP6 or HBM which when expressed in a cell is capable of modulating bone mass, lipid levels, Dkk activity, and/or Wnt activity.

In one embodiment of the present invention, "*HBM* gene" refers to the genomic DNA sequence found in individuals showing the HBM characteristic or phenotype, where the sequence encodes the protein indicated by SEQ ID NO: 4. The *HBM* gene and the *Zmax1* (*LRP5*) gene are allelic. The protein encoded by the *HBM* gene has the property of causing elevated bone mass, while the protein encoded by the *Zmax1* (*LRP5*) gene does not. The *HBM* gene and the *Zmax1* (*LRP5*) gene differ in that the *HBM* gene has a thymine at position 582, while the *Zmax1* gene has a guanine at position 582. The *HBM* gene comprises the nucleic acid sequence shown as SEQ ID NO: 2. The *HBM* gene may also be referred to as an "HBM polymorphism." Other *HBM* genes may further have silent mutations, such as those discussed in Section 3 below.

In alternative embodiments of the present invention, "HBM gene" may also refer to any allelic variant of Zmax1 (LRP5) or LRP6 which results in the HBM phenotype. Such variants may include alteration from the wild-type protein coding sequence as described herein and/or alteration in expression control sequences of Zmax1 (LRP5) or contains an amino acid mutation in LRP5 or LRP6, such that the resulting protein produces a phenotype which enhances bone mass and/or modulates lipid levels. A preferred example of such a variant is an alteration of the endogenous Zmax1 (LRP5) promoter region resulting in increased expression of the Zmax1 (LRP5) protein.

"Normal," "wild-type," "unaffected", "Zmax1", "Zmax", "LR3" and "LRP5" all refer to the genomic DNA sequence that encodes the protein indicated by SEQ ID NO: 3. LRP5 has also been referred to LRP7 in mouse. Zmax1, LRP5 and Zmax may be used interchangeably throughout the specification and are meant to be the same gene, perhaps only relating to the gene in a different organism. The *Zmax1* gene has a guanine at position 582 in the human sequence. The *Zmax1* gene of human comprises the nucleic acid sequence shown as SEQ ID NO: 1. "Normal," "wild-type," "unaffected", "Zmax1" and "LRP5" also refer to allelic variants of the genomic sequence that encodes proteins that do not contribute to elevated bone

mass. The *Zmax1* (*LRP5*) gene is common in the human population, while the *HBM* gene is rare.

"Bone development" generally refers to any process involved in the change of bone over time, including, for example, normal development, changes that occur during disease states, and changes that occur during aging. This may refer to structural changes and dynamic rate changes such as growth rates, resorption rates, bone repair rates, and etc. "Bone development disorder" particularly refers to any disorders in bone development including, for example, changes that occur during disease states and changes that occur during aging. Bone development may be progressive or cyclical in nature. Aspects of bone that may change during development include, for example, mineralization, formation of specific anatomical features, and relative or absolute numbers of various cell types.

"Bone modulation" or "modulation of bone formation" refers to the ability to affect any of the physiological processes involved in bone remodeling, as will be appreciated by one skilled in the art, including, for example, bone resorption and appositional bone growth, by, *inter alia*, osteoclastic and osteoblastic activity, and may comprise some or all of bone formation and development as used herein.

Bone is a dynamic tissue that is continually adapting and renewing itself through the renewal of old or unnecessary bone by osteoclasts and the rebuilding of new bone by osteoblasts. The nature of the coupling between these processes is responsible for both the modeling of bone during growth as well as the maintenance of adult skeletal integrity through remodeling and repair to meet the everyday needs of mechanical usage. There are a number of diseases that result from an uncoupling of the balance between bone resorption and formation. With aging there is a gradual "physiologic" imbalance in bone turnover, which is particularly exacerbated in women due to menopausal loss of estrogen support, that leads to a progressive loss of bone. As bone mineral density falls below population norms there is a consequent increase in bone fragility and susceptibility to spontaneous fractures. For every 10 percent of bone that is lost, the risk of fracture doubles. Individuals with bone mineral density (BMD) in the spine or proximal femur 2.5 or

more standard deviations below normal peak bone mass are classified as osteoporotic. However, osteopenic individuals with BMD between 1 and 2.5 standard deviations below the norm are clearly at risk.

Bone is measured by several different forms of X-ray absorptiometry. All of the instruments measure the inorganic or bone mineral content of the bone. Standard DXA measurements give a value that is an areal density, not a true density measurement by the classical definition of density (mass/unit volume). Nevertheless, this is the type of measurement used clinically to diagnose osteoporosis. However, while BMD is a major contributing factor to bone strength, as much as 40% of bone strength stems from other factors including: 1) bone size (*i.e.*, larger diameters increase organ-level stiffness, even in the face of lower density); 2) the connectivity of trabecular structures; 3) the level of remodeling (remodeling loci are local concentrators of strain); and 4) the intrinsic strength of the bony material itself, which in turn is a function of loading history (*i.e.*, through accumulated fatigue damage) and the extent of collagen cross-linking and level of mineralization. There is good evidence that all of these strength/fragility factors play some role in osteoporotic fractures, as do a host of extraskeletal influences as well (such as fall patterns, soft tissue padding, and central nervous system reflex responsiveness).

Additional analytical instruments can be used to address these features of bone. For example, the pQCT allows measurement of separate trabecular and cortical compartments for size and density and the μCT provides quantitative information on architectural features such as trabecular connectivity. The μCT also gives a true bone density measurement. With these tools, the important non-BMD parameters can be measured for diagnosing the extent of disease and the efficacy of treatments. Current treatments for osteoporosis are based on the ability of drugs to prevent or retard bone resorption. Although newer anti-resorptive agents are proving to be useful in the therapy of osteoporosis, they are viewed as short-term solutions to the more definitive challenge to develop treatments that will increase bone mass and/or the bone quality parameters mentioned above.

Thus, bone modulation may be assessed by measuring parameters such as bone mineral density (BMD) and bone mineral content (BMC) by pDXA X-ray methods, bone size, thickness or volume as measured by X-ray, bone formation rates as measured for example by calcien labeling, total, trabecular, and mid-shaft density as measured by pQCT and/or μCT methods, connectivity and other histological parameters as measured by μCT methods, mechanical bending and compressive strengths as preferably measured in femur and vertebrae respectively. Due to the nature of these measurements, each may be more or less appropriate for a given situation as the skilled practitioner will appreciate. Furthermore, parameters and methodologies such as a clinical history of freedom from fracture, bone shape, bone morphology, connectivity, normal histology, fracture repair rates, and other bone quality parameters are known and used in the art. Most preferably, bone quality may be assessed by the compressive strength of vertebra when such a measurement is appropriate. Bone modulation may also be assessed by rates of change in the various parameters. Most preferably, bone modulation is assessed at more than one age.

"Normal bone density" refers to a bone density within two standard deviations of a Z score of 0 in the context of the HBM linkage study. In a general context, the range of normal bone density parameters is determined by routine statistical methods. A normal parameter is within about 1 or 2 standard deviations of the age and sex normalized parameter, preferably about 2 standard deviations. A statistical measure of meaningfulness is the P value which can represent the likelihood that the associated measurement is significantly different from the mean. Significant P values are $P < 0.05$, 0.01, 0.005, and 0.001, preferably at least $P < 0.01$.

"HBM" refers to "high bone mass" although this term may also be expressed in terms of bone density, mineral content, and size.

The "HBM phenotype" and "HBM-like phenotype" may be characterized by an increase of about 2 or more standard deviations, preferably 2, 2.5, 3, or more standard deviations in 1, 2, 3, 4, 5, or more quantitative parameters of bone modulation, preferably bone density and mineral content and bone strength

parameters, above the age and sex norm for that parameter. The HBM phenotype and HBM-like phenotype are characterized by statistically significant increases in at least one parameter, preferably at least 2 parameters, and more preferably at least 3 or more parameters. The HBM phenotype and the HBM-like phenotype may also be characterized by an increase in one or more bone quality parameters and most preferably increasing parameters are not accompanied by a decrease in any bone quality parameters. Most preferably, an increase in bone modulation parameters and/or bone quality measurements is observed at more than one age. The HBM phenotype and HBM-like phenotype also includes changes of lipid levels, Wnt activity and/or Dkk activity.

The terms "isolated" and "purified" refer to a substance altered by hand of man from the natural environment. An isolated peptide may be for example in a substantially pure form or otherwise displaced from its native environment such as by expression in an isolated cell line or transgenic animal. An isolated sequence may for example be a molecule in substantially pure form or displaced from its native environment such that at least one end of said isolated sequence is not contiguous with the sequence it would be contiguous with in nature.

"Biologically active" refers to those forms of proteins and polypeptides, including conservatively substituted variants, alleles of genes encoding a protein or polypeptide fragments of proteins which retain a biological and/or immunological activity of the wild-type protein or polypeptide. Preferably the activity is one which induces a change in Dkk activity, such as inhibiting the interaction of Dkk with a ligand binding partner (*e.g.,* LRP5 or LRP6 or Dkk-1 with a Dkk-1 interacting protein such as those shown in Figure 5). By biologically active is also meant to include any form which modulates Wnt signaling.

By "modulate" and "regulate" is meant methods, conditions, or agents which increase or decrease the wild-type activity of an enzyme, inhibitor, signal transducer, receptor, transcription activator, co-factor, and the like. This change in activity can be an increase or decrease of mRNA translation, mRNA or DNA transcription, and/or

mRNA or protein degradation, which may in turn correspond to an increase or decrease in biological activity.

By " modulated activity" is meant any activity, condition, disease or phenotype which is modulated by a biologically active form of a protein. Modulation may be effected by affecting the concentration or subcellular localization of biologically active protein, *i.e.*, by regulating expression or degradation, or by direct agonistic or antagonistic effect as, for example, through inhibition, activation, binding, or release of substrate, modification either chemically or structurally, or by direct or indirect interaction which may involve additional factors.

By "effective amount" or "dose effective amount" or "therapeutically effective amount" is meant an amount of an agent which modulates a biological activity of the polypeptide of the invention.

By "immunologically active" is meant any immunoglobulin protein or fragment thereof which recognizes and binds to an antigen.

By "Dkk" is meant to refer to the nucleic acids and proteins of members of the Dkk (Dickkopf) family. This includes, but is not limited to, Dkk-1, Dkk-2, Dkk-3, Dkk-4, Soggy, and related Dkk proteins. Dkk-1 is a preferred embodiment of the present invention. However, the Dkk proteins have substantial homology and one skilled in the art will appreciate that all of the embodiments of the present invention utilizing Dkk-1 may also be utilized with the other Dkk proteins.

By "Dkk-1" is meant to refer to the Dkk-1 protein and nucleic acids which encode the Dkk-1 protein. Dkk-1 refers to Dickkopf-1, and in *Xenopus* it is related to at least Dkk-2, Dkk-3, and Dkk-4 (*see* Krupnik *et al.*, *Gene* 238:301-313 (1999)). Dkk-1 was first identified in *Xenopus* (Glinka *et al.*, *Nature* 391:357-62 (1998)). It was recognized as a factor capable of inducing ectopic head formation in the presence of inhibition of the BMP pathway. It was then also found to inhibit the axis-inducing activity of several *Xenopus* Wnt molecules by acting as an extracellular antagonist of Wnt signaling. Mammalian homologs have been found including Dkk-1, Dkk-2, Dkk-3, Dkk-4 and soggy (Fedi *et al.*, 1999 and Krupnick *et al.* 1999). Human Dkk-1 was also referred to as sk (Fedi *et al.* 1999). As used herein, Dkk-1 is

meant to include proteins from any species having a Wnt pathway in which Dkk-1 interacts. Particularly preferred are mammalian species (*e.g.*, murine, caprine, canine, bovine, feline, equine, primate, ovine, porcine and the like), with particularly preferred mammals being humans. Nucleic acid sequences encoding Dkk-1 include, but are not limited to human Dkk-1 (GenBank Accession Nos. AH009834, XM_005730, AF261158, AF261157, AF177394, AF127563 and NM_012242), *Mus musculus* dickkopf homolog 1 (GenBank Accession No. NM_010051), and *Danio rerio* dickkopf-1 (GenBank Accession Nos. AF116852 and AB023488). The genomic sequences with exon annotation are GenBank Accession Nos. AF261157 and AF261158. Also contemplated are homologs of these sequences which have Dkk-1 activity in the Wnt pathway. Dkk-1 amino acid sequences include, but are not limited to human dickkopf homolog 1 (GenBank Accession Nos. AAG15544, BAA34651, NP_036374, AAF02674, AAD21087, and XP_005730), *Danio rerio* (zebrafish) dickkopf1 (GenBank Accession Nos. BAA82135 and AAD22461) and murine dickkopf-1 (GenBank Accession Nos. O54908 and NP_034181). Variants and homologs of these sequences which possess Dkk-1 activity are also included when referring to Dkk-1.

By "Dkk mediated" disorder, condition or disease is any abnormal state that involves Dkk activity. The abnormal state can be induced by environmental exposure or drug administration. Alternatively, the disease or disorder can be due to a genetic defect. Dkk mediated diseases, disorders and conditions include but are not limited to bone mass disorders or conditions and lipid disorders and conditions. For example, bone mass disorders/conditions/diseases, which may be mediated by Dkk, include but are not limited to age related loss of bone, bone fractures (*e.g.,* hip fracture, Colle's fracture, vertebral crush fractures), chondrodystrophies, drug-induced disorders (e.g., osteoporosis due to administration of glucocorticoids or heparin and osteomalacia due to administration of aluminum hydroxide, anticonvulsants, or glutethimide), high bone turnover, hypercalcemia, hyperostosis, osteogenesis imperfecta, osteomalacia, osteomyelitis, osteoporosis, Paget's disease, osteoarthritis, and rickets.

Lipid disorders/diseases/conditions, which may be mediated by Dkk, include but are not limited to familial lipoprotein lipase deficiency, familial apoprotein CII deficiency, familial type 3 hyperlipoproteinemia, familial hypercholesterolemia, familial hypertriglyceridemia, multiple lipoprotein-type hyperlipidemia, elevated lipid levels due to dialysis and/or diabetes, and elevated lipid levels of unknown etiologies

The term "recognizes and binds," when used to define interactions of antisense nucleotides, siRNAs (small inhibitory RNA),·or shRNA (short hairpin RNA) with a target sequence, means that a particular antisense, siRNA, or shRNA sequence is substantially complementary to the target sequence, and thus will specifically bind to a portion of an mRNA encoding polypeptide. As such, typically the sequences will be highly complementary to the mRNA target sequence, and will have no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 base mismatches throughout the sequence. In many instances, it may be desirable for the sequences to be exact matches, i.e. be completely complementary to the sequence to which the oligonucleotide specifically binds, and therefore have zero mismatches along the complementary stretch. As such, highly complementary sequences will typically bind quite specifically to the target sequence region of the mRNA and will therefore be highly efficient in reducing, and/or even inhibiting the translation of the target mRNA sequence into polypeptide product.

Substantially complementary oligonucleotide sequences will be greater than about 80 percent complementary (or `% exact-match`) to the corresponding mRNA target sequence to which the oligonucleotide specifically binds, and will, more preferably be greater than about 85 percent complementary to the corresponding mRNA target sequence to which the oligonucleotide specifically binds. In certain aspects, as described above, it will be desirable to have even more substantially complementary oligonucleotide sequences for use in the practice of the invention, and in such instances, the oligonucleotide sequences will be greater than about 90 percent complementary to the corresponding mRNA target sequence to which the oligonucleotide specifically binds, and may in certain embodiments be greater than about 95 percent complementary to the corresponding mRNA target sequence to

which the oligonucleotide specifically binds, and even up to and including 96%, 97%, 98%, 99%, and even 100% exact match complementary to the target mRNA to which the designed oligonucleotide specifically binds.

Percent similarity or percent complementary of any of the disclosed sequences may be determined, for example, by comparing sequence information using the GAP computer program, version 6.0, available from the University of Wisconsin Genetics Computer Group (UWGCG). The GAP program utilizes the alignment method of Needleman and Wunsch (1970). Briefly, the GAP program defines similarity as the number of aligned symbols (i.e., nucleotides or amino acids) which are similar, divided by the total number of symbols in the shorter of the two sequences. The preferred default parameters for the GAP program include: (1) a unary comparison matrix (containing a value of 1 for identities and 0 for non-identities) for nucleotides, and the weighted comparison matrix of Gribskov and Burgess (1986), (2) a penalty of 3.0 for each gap and an additional 0.10 penalty for each symbol in each gap; and (3) no penalty for end gaps.

By "mimetic" is meant a compound or molecule that performs the same function or behaves similarly to the compound mimicked.

By "reporter element" is meant a polynucleotide that encodes a poplypeptide capable of being detected in a screening assays. Examples of polypeptides encoded by reporter elements include, but are not limited to, lacZ, GFP, luciferase, and chloramphenicol acetyltransferase.

2. Introduction

A polymorphism in LRP5 (Zmax), G171V, designated as HBM, has been identified as conferring a high bone mass phenotype in a population of related subjects as described in co-pending applications International Patent Application PCT/US 00/16951, and U.S. Patent Application Nos. 09/543,771 and 09/544,398, which are hereby incorporated by reference in their entirety (Little *et al., Am J Hum Genet.* 70:11-19 (2002)). LRP5 is also described in International Patent Application WO 98/46743, which is incorporated by reference in its entirety. Loss of LRP5

function has been shown to have a deleterious effect on bone (Gong *et al., Cell* 107:513-523 (2001)). Additionally, the HBM polymorphism and LRP5 may also be important in cardiac health and lipid-mediated disorders. Thus, methods of regulating their activity can serve as methods of treating and/or preventing cardiac and lipid-mediated disorders.

Recent studies have indicated that LRP5 participates in the Wnt signal transduction pathway. The Wnt pathway is critical in limb early embryological development. A recently published sketch of the components of Wnt signaling is shown in Figure 1 (Nusse, 2001 http://www.stanford.edu/~rnusse/pathways/cell2.html) (*see also*, Nusse, *Nature* 411:255-6 (2001); and Mao *et al., Nature* 411:321-5 (2001)). Briefly summarized, Wnt proteins are secreted proteins which interact with the transmembrane protein Frizzled (Fz). LRP proteins, such as LRP5 and LRP6, are believed to modulate the Wnt signal in a complex with Fz (Tamai *et al., Nature* 407:530-5 (2000)). The Wnt pathway acts intracellularly through the Disheveled protein (Dsh) which in turn inhibits glycogen synthetase kinase-3 (GSK3) from phosphorylating β-catenin. Phosphorylated β-catenin is rapidly degraded following ubiquitination. However, the stabilized β-catenin accumulates and translocates to the nucleus where it acts as a cofactor of the T-cell factor (TCF) transcription activator complex.

The protein dickkopf-1 (Dkk-1) is reported to be an antagonist of Wnt pathway. Dkk-1 is required for head formation in early development. Dkk-1 and its function in the Wnt pathway are described in *e.g.,* Krupnik, *et al., Gene* 238:301-13 (1999); Fedi *et al., J. Biol. Chem.* 274:19465-72 (1999); *see also* for Dkk-1 and the Wnt pathway, Wu *et al., Curr. Biol.* 10:1611-4 (2000), Shinya *et al., Mech. Dev.* 98:3-17 (2000), Mukhopadhyay *et al., Dev Cell* 1:423-434 (2001) and in PCT Patent Application No. WO 00/52047, and in references cited in each. It has been known that Dkk-1 acts upstream of Dsh, however the nature of the mechanism of inhibition by Dkk-1 is just beginning to be elucidated. Dkk-1 is expressed in the mouse embryonic limb bud and its disruption results in abnormal limb morphogensis, among

other developmental defects (Gotewold *et al., Mech. Dev.* 89:151-3 (1999); and, Mukhopadhyay *et al., Dev Cell* 1:423-434 (2001)).

Related U.S. provisional application 60/291,311 disclosed a novel interaction between Dkk-1 (GenBank Accession No. XM 005730) and LRP5. The interaction between Dkk-1 and LRP5 was discovered by a yeast two hybrid (Y2H) screen for proteins which interact with the ligand binding domain of LRP5, as described in Example 1. The two-hybrid screen is a common procedure in the art, which is described, for example, by Gietz *et al., Mol. Cell. Biochem.* 172:67-79 (1997); Young, *Biol. Reprod.* 58:302-11 (1998); Brent and Finley, *Ann. Rev. Genet.* 31:663-704 (1997); and Lu and Hannon, eds., Yeast Hybrid Technologies, Eaton Publishing, Natick MA, (2000). More recently, other studies confirm that Dkk-1 is a binding partner for LRP and modulates the Wnt pathway via direct binding with LRP (R. Nusse, *Nature* 411:255-256 (2001); A. Bafico *et al., Nat. Cell Biol.* 3:683-686 (2001); M. Semënov, *Curr. Biol.* 11:951-961 (2001); B. Mao, *Nature* 411:321-325 (2001), Zorn, *Curr. Biol.* 11:R592-5 (2001)); and, L. Li *et al., J. Biol Chem.* 277:5977-81 (2002)).

Mao and colleagues (2001) identified Dkk-1 as a ligand for LRP6. Mao *et al.* suggest that Dkk-1 and LRP6 interact antagonistically where Dkk proteins inhibit the Wnt coreceptor functions of LRP6. Using co-immunoprecipitation, the group verified that the Dkk-1/LRP6 interaction was direct. Dkk-2 was also found to directly bind LRP6. Contrary to data contained in provisional application 60/291,311, Mao *et al.* report that no interaction was detected between any Dkk protein and LRP5, as well as no interaction with LDLR, VLDLR, ApoER, or LRP). Additionally, Mao *et al.* demonstrated that LRP6 can titrate Dkk-1's effects of inhibiting Wnt signaling using the commercial TCF-luciferase reporter gene assay (TOPFLASH). A similar conclusion was drawn from analogous studies in *Xenopus* embryos. Deletion analyses of LRP6 functional domains revealed that EGF repeats (beta-propellers) 3 and 4 were necessary for Dkk-1 binding and that the ligand binding domains of LRP6 had no effect on Dkk-1 binding. The findings of Mao *et al.* contrast with data obtained by the present inventors indication that the ligand binding domains of LRP5

were necessary and sufficient for Dkk-1 binding in yeast. Using classical biochemical ligand-receptor studies, Mao *et al.* determined a Kd=0.34 nM for Dkk-1/LRP6 and a Kd=0.73 nM for Dkk-2/LRP6.

Semenov *et al.* (2001) verified the Mao group's results and confirmed by coimmunoprecipitation that Dkk-1 does not directly bind to Wnt or Frizzled but rather interacts with LRP6. Their Scatchard analyses found a Kd=0.5 nM for Dkk-1/LRP6. Semenov *et al.* also demonstrated that Dkk-1 could abolish an LRP5/Frizzled8 complex implying that Dkk-1 can also repress Wnt signaling via interactions with LRP5. A Dkk-1 mutant where cysteine 220 was changed to alanine abolished LRP6 binding and was unable to repress Wnt signaling. Studies in *Xenopus* embryos confirmed the results and revealed a functional consequence of Dkk-1/LRP6: repression of Wnt signaling. Their *Xenopus* work also suggested that LRP6/Dkk-1 may be specific for the canonical, β-catenin-mediated, Wnt pathways as opposed to the Wnt Planar Cell Polarity pathway.

Bafico *et al.* (2001) employed a ^{125}I-labeled Dkk-1 molecule to identify LRP6 as its sole membrane receptor with a Kd=0.39 nM. Again, the functional consequences of the Dkk-1/LRP6 interaction was a repression of the canonical Wnt signaling even when Dkk-1 was added at extremely low concentrations (30 pM).

Not wishing to be bound by theory, it is believed that the present invention provides an explanation for the mechanism of Dkk-1 inhibition of the Wnt pathway and provides a mechanism whereby the Wnt pathway may be modulated. The present application and related provisional application 60/291,311 describe Dkk-1/LRP5 interactions and demonstrate that the interaction between LRP5/LRP6/HBM and Dkk can be used in a method as an intervention point in the Wnt pathway for an anabolic bone therapeutic or a modulator of lipid metabolism.

As detailed below, in the section "Methods to Identify Binding Partners" and Examples 6 and 7, Dkk-1 is able to repress LRP5-mediated Wnt signaling but not HBM-mediated Wnt signaling. This observation is of particular interest because the HBM mutation in LRP5 is a gain of function or activation mutation. That is, Wnt signaling, via the canonical pathway, is enhanced with HBM versus LRP5. The

present data suggest the mechanism of this functional activation: the inability of Dkk-1 to repress HBM-mediated Wnt signaling. Further investigations of other Wnt or Dkk family members show differential activities in the canonical Wnt pathway that demonstrate the complexity and variability in Wnt signaling that can be achieved depending on the LRP/Dkk/Wnt/Frizzled repertoire that is expressed in a particular cell or tissue. This may attest to the apparent bone specificity of the HBM phenotype in humans and in the HBM transgenic animals.

Furthermore, the present data reveal the importance and functional consequence for the potential structural perturbation of the first beta-propeller domain of LRP5. Our data identified the ligand binding domain of LRP5 as the interacting region with Dkk-1 while the Mao *et al.* publication demonstrated the functional role of propellers 3 and 4 in their LRP6/Dkk-1 studies. In the present invention, we implicate the first beta propeller domain, via the HBM mutation at residue 171, as having a functional consequence in the Dkk-1-mediated Wnt pathway. The involvement of position 171 of propeller 1 may be direct or indirect with Dkk-1. Direct involvement could arise from perturbations of the 3-dimensional structure of the HBM extracellular domain that render Dkk-1 unable to bind. Alternatively, residue 171 of propeller 1 may directly interact with Dkk-1; however, by itself, it is insufficient to bind and requires other LRP5 domains. Potential indirect candidate molecules may be among the proteins identified the Dkk-1 yeast-two-hybrid experiments.

It may be that the disruption of Dkk activity is not necessarily mediated by enhancing or preventing the binding of Dkk to LRP5/LRP6/HBM. More than one mechanism may be involved. Indeed, the inventors have observed that Dkk-1 binds LRP5, LRP6, and HBM. It is able to effectively inhibit LRP6, and to a slightly lesser extent, LRP5 activity. Further, has been observed that different members of the Dkk family differentially affect LRP5/LRP6/HBM activity. For example, Dkk-1 inhibits LRP5/LRP6/HBM activity while another Dkk may enhance LRP5/LRP6/HBM activity. An endpoint to consider is the modulation of the LRP5/LRP6/HBM activity, not simply binding.

The present disclosure shows that targeting the disruption of the Dkk-1/LRP5 interaction is a therapeutic intervention point for an HBM mimetic agent. A therapeutic agent of the invention may be a small molecule, peptide or nucleic acid aptamer, antibody, or other peptide/protein, etc. Methods of reducing Dkk-1 expression may also be therapeutic using methodologies such as: RNA interference, antisense oligonucleotides, morpholino oligonucleotides, PNAs, antibodies to Dkk-1 or Dkk-1 interacting proteins, decoy or scavenger LRP5 or LRP6 receptors, and knockdown of Dkk-1 or Dkk-1 interactor transcription.

In an embodiment of the present invention, the activity of Dkk-1 or the activity of a Dkk-1 interacting protein may be modulated for example by binding with a peptide aptamer of the present invention. In another embodiment, LRP5 activity may be modulated by a reagent provided by the present invention (*e.g.,* a peptide aptamer). In another embodiment, the Dkk-1/LRP5 interaction may be modulated by a reagent of the present invention (*e.g.*, a Dkk-1 interacting protein such as those identified in Figure 5). In another embodiment, the Wnt signal transduction pathway may be modulated by use of one or more of the above methods. In a preferred embodiment of the present invention, the Dkk-1 mediated activity of the Wnt pathway may be specifically modulated by one or more of the above methods. In another preferred embodiment of the present invention, the Wnt signal transduction pathway may be stimulated by down-regulating Dkk-1 interacting protein activity; such down-regulation could, for example, yield greater LRP5 activity. In a more preferred embodiment, by stimulating LRP5 activity, bone mass regulation may be stimulated to restore or maintain a more optimal level. In another preferred embodiment, by stimulating LRP5 activity, lipid metabolism may be stimulated to restore or maintain a more optimal level. Alternative embodiments provide methods for screening candidate drugs and therapies directed to correction of bone mass disorders or lipid metabolism disorders. And, preferred embodiments of the present invention provide drugs and therapies developed by the use of the reagents and/or methods of the present invention. One skilled in the art will understand that the present invention provides important research tools to develop an effective model of

osteoporosis, to increase understanding of bone mass and lipid modulation, and to modulate bone mass and lipid metabolism.

Previous investigation of a large family in which high bone mass is inherited as a single gene (autosomal dominant) trait (HBM-1) has provided important insight into the mechanism by which bone density might be modulated. Members of this family have significantly increased spinal and hip BMD (>3 standard deviations above the norm) which affects young adults as well as elderly family members into the ninth decade. The bones of affected members, while appearing very dense radiographically, have normal external shape and outer dimensions. Cortical bone is thickened on endosteal surfaces and "affected" individuals are asymptomatic without any other phenotypic abnormalities. Assays of biochemical markers that reflect skeletal turnover suggest that the disorder is associated with a normal rate of bone remodeling. Affected individuals have achieved a balance in bone turnover at a density that is significantly greater than necessary for normal skeletal stresses. Importantly, the bones most affected are load-bearing bones which are subjected to the greatest mechanical and gravitational stresses (spine and hip). These are the most important bones to target fir therapeutic interventions in osteoporosis. The gene identified as being responsible for this phenotype, Zmax or LRP5, was not previously associated with bone physiology. The fact that modification of this gene, such as that produced by the polymorphism leading to the autosomal dominant inheritance of the HBM family phenotype, identifies Zmax/LRP5 and the pathway by which it is regulated, including Dkk/Wnt pathways discussed above, as an important target for developing modulators of bone density. Modulation of Zmax/LRP5 to mimic the gain in function provided by the HBM polymorphism would be expected to provide an important therapy for bone wasting conditions. Additionally, such modulation in young adults could enhance peak bone mass and prevent or delay fracture risk later in life. Alternatively, modulation to reduce function could be employed to treat conditions where bone is being inappropriately produced.

3. Polypeptides

Polypeptides contemplated for use in this invention include those which modulate Dkk and Dkk interacting protein activities. Preferred polypeptides and peptides include those which modulate the Wnt pathway. Examples of preferred sequences include the Y2H baits exemplified in Figure 2, peptide aptamers of Figure 3 (SEQ ID NOs:171-188) and Figure 4 (SEQ ID NOs:189-192), the polypeptides of the Dkk-1 interacting proteins identified in Figure 5, those polypeptides shown in Figure 6, the LRP binding domain of Dkk (amino acids 138-266 of hDkk1), the cysteine-rich domain 2 (a.a. 183-245 of hDkk-1), the cysteine-rich domain 1 (a.a. 97-138 of hDkk), and LRP5 binding aptamers of Figure 13 (including SEQ ID NOs:204-213). Although Dkk-1 is exemplified, the other Dkk proteins contain substantially similar regions and may also be used according to the present invention.

For example, the baits depicted in Figure 2 were used in a yeast two hybrid (Y2H) screen. The Y2H screen was performed as described in Example 2 to determine the minimum required binding domain for Dkk-1 to bind LRP5. The minimum binding domain constructs (*i.e.,* residues 139-266 in bold below and residues 97-245 which are underlined, of Dkk-1) include the second cysteine rich domain which has sequence homology to a colipase fold.

mmalgaagat rvfvamvaaa lgghpllgvs atlnsvlnsn aiknlppplg gaaghpgsav 60
saapgilypg gnkyqtidny qpypcaedee cgtdeycasp trggdagvqi clacrkrrkr 120
cmrhamccpg nycknqicvs sdqnhfrqei eetitesfgn dhstldgysr rttlsskmyh 180
tkgqegsvcl rssdcasglc carhfwskic kpvlkegqvc tkhrrkgshg leifqrcycg 240
eglscriqkd hhqasnssrl htcqrh (GenBank Accession No. XP_005730) (SEQ ID NO:128).

This homology suggests a lipid-binding function and may facilitate Dkk-1 interactions at the plasma membrane (van Tilbeurgh, H., *Biochim. Biophys. Acta.* 1441:173-84 (1999)). An interaction domain of Dkk-1 that is able to interact with the ligand binding domain (LBD) of LRP5 is a useful reagent in the modulation of LRP5 activity

and modulation of Dkk-1/LRP5 complex formation. Similar screens can be prepared for Dkk-1 and Dkk-1 interacting proteins or polypeptides.

A set of peptide aptamers was identified from a library of random peptides constrained and presented in a thioredoxin A (trxA) scaffold as described in Example 3. Peptide aptamers are powerful new tools for molecular medicine as reviewed by Hoppe-Seyler & Butz, *J. Mol. Med.*, 78:426-430 (2000); Brody and Gold, *Rev. Mol. Biotech.*, 74:5-13 (2000); and Colas, *Curr. Opin. in Chem. Biol.* 4:54-9 (2000) and the references cited therein. Briefly, peptide aptamers have been shown to be highly specific reagents capable of binding *in vivo*. As such, peptide aptamers provide a method of modulating the function of a protein and may serve as a substitute for conventional knock-out methods, knock-down or complete loss of function. Peptide aptamers are also useful reagents for the validation of targets for drug development and may be used as therapeutic compounds directly or provide the necessary foundation for drug design. Once identified, the peptide insert may be synthesized and used directly or incorporated into another carrier molecule. References reviewed and cited by Brody and Gold (2000, *supra*) describe demonstrated therapeutic and diagnostic applications of peptide aptamers and would be known to the skilled artisan.

The peptide aptamers of the present invention are useful reagents in the binding of Dkk-1 to its ligands and thereby modulation of the Wnt pathway and may be used to prevent Dkk-1 from inhibiting LRP5 modulation or Dkk-1 interacting protein modulation of the Wnt pathway. The sequence of these peptide aptamers is shown in Figure 3 (SEQ ID NOs:171-188). The peptide aptamers refers to the peptide constrained by the thioredoxin scaffold. The aptamers are also contemplated as therapeutic agents to treat Dkk-1 mediated diseases and conditions. Such aptamers are useful structural guides to chemists, for the design of mimetic compounds of the aptamers.

Peptide aptamers were likewise developed to the LRP5 ligand binding domain (LBD) bait sequences. The sequences of these peptide aptamers is shown in Figure 4 (SEQ ID NOs:189-192). These are useful reagents which may be used to disrupt

the Dkk-1/LRP5 binding interface while leaving Dkk-1 undisturbed. These can be used as comparative controls for Wnt signaling, thus, a control is provided for the specificity of any drug or therapy screened. The aptamers are also useful therapeutic agents to treat LRP mediated diseases and conditions. Such aptamers may also be used as structural guides to chemists, for the design of mimetic compounds of the aptamers.

Thirty proteins were identified which interact with Dkk-1, Dkk-1 interacting proteins, were identified in a yeast-two-hybrid screen using the Dkk-1 bait and are shown in Figure 5. It was noted that these results suggest an interaction of Dkk-1 with Notch-2. It has been suggested that cross-talk exists between the Wnt and Notch signaling pathways. For instance, Presenilin1 (Ps1) is required for Notch processing and inhibits the downstream Wnt pathway. The extracellular domain of Notch is thought to interact with Wnt. Furthermore, the Notch intracellular domain is thought to interact with disheveled and in signal induced processing, the intracellular domain is thought to interact with presenilin. (Soriano *et al.*, *J. Cell Biol.* 152:785-94 (2001)). For additional information regarding the relationships between Notch and Wnt signaling, see Wesley, *Mol. Cell. Biol.* 19:5743-58 (1999) and Axelrod *et al.*, *Science* 271:1826-32 (1996).

An interaction between Dkk-1 and chordin has also been noted; suggesting that cross-talk exists between the Wnt and TGF-beta/BMP signaling pathways (Letamendia *et al., J. Bone Joint Surg. Am.* 83A:S31 (2001); Labbe *et al., Proc. Natl. Acad. Sci. USA* 97:8358-63 (2000); Nishita *et al., Nature* 403:781-5 (2000); DeRobertis *et al., Int. J. Dev. Biol.* 45:1389-97 (2001); and Saint-Jeannet *et al., Proc. Natl. Acad. Sci. USA* 94:13713-8 (1997)). The BMP signaling pathway has an established role in bone and connective tissue development, repair and homeostasis (review in Rosen and Wozney "Bone Morphogenetic Proteins" In: Principles of Bone Biology, 2nd Edition, Eds. J. Bilezikian, L. Raisz and G. Rodan, Academic Press, pp. 919-28 (2002)). Chordin is an important molecule during development which also modulates BMP signaling in adults by sequestering BMPs in latent complexes (Piccolo *et al., Cell* 86:589-98 (1996) reviewed in Reddi, *Arthritis Res.* 3:1-5 (2001);

DeRobertis *et al., Int. J. Dev. Biol.* 45:189-97 (2001)). It may be that Dkk effects bone mass modulation through both the Wnt signaling pathway via LRP and the BMP pathway via chordin.

Moreover, a number of putative growth factors, growth factor related proteins, and extracellular matrix proteins have been identified as Dkk-1 interacting proteins. Additional information regarding Dkk-1 interacting proteins identified in the Y2H assay may be obtained from publicly available databases such as PubMed via the use of the accession numbers provided in the present application. In a preferred embodiment of the invention, the amino acid sequences of these Dkk-1 interacting proteins or biologically active fragments thereof be used to modulate Dkk, Dkk-1, LRP5, LRP6, HBM, or Wnt activity. Although these proteins were identified as interacting with Dkk-1, due to the substantial homology between the various Dkk proteins, such interacting proteins are contemplated to interact with the other Dkk family members.

4. Aptamer Mimetics

The present invention further provides for mimetics of Dkk, particularly Dkk-1, and LRP5 peptide aptamers. Such aptamers may serve as structural guides to chemists for the design of mimetic compounds of the aptamers. The aptamers and their mimetics are useful as therapeutic agents to treat LRP- or Dkk-mediated diseases and conditions.

5. Nucleic Acid Molecules

The present invention further provides nucleic acid molecules that encode polypeptides and proteins which interact with Dkk and Dkk interacting proteins, and/or LRP5 (also LRP6 and HBM) to modulate biological activities of these proteins. Preferred embodiments provide nucleic acids encoding for fragments of Dkk-1 protein, including the nucleic acids of Figure 7, the Dkk-1 interacting proteins listed in Figure 5, polypeptide aptamers of Dkk-1 (Figure 3 - SEQ ID NOs:171-188), LRP5 (Figure 4 - SEQ ID NOs:189-192), Figure 13 peptide aptamers (including SEQ

ID NO:204-214) encoded by Figure 12 polynucleotides (including SEQ ID NO:193-203), LRP6 and HBM and the related fusion proteins herein described, preferably in isolated or purified form. As used herein, "nucleic acid" is defined as RNA, DNA, or cDNA that encodes a peptide as defined above, or is complementary to a nucleic acid sequence encoding such peptides, or hybridizes to either the sense or antisense strands of the nucleic acid and remains stably bound to it under appropriate stringency conditions. The nucleic acid may encode a polypeptide sharing at least about 75% sequence identity, preferably at least about 80%, and more preferably at least about 85%, with the peptide sequences; at least about 90%, 95%, 96%, 97%, 98%, and 99% or greater are also contemplated. Specifically contemplated are genomic DNA, cDNA, mRNA, antisense molecules, enzymatically active nucleic acids (*e.g.,* ribozymes), as well as nucleic acids based on an alternative backbone or including alternative bases, whether derived from natural sources or synthesized. Such hybridizing or complementary nucleic acids, however, are defined further as being novel and nonobvious over any prior art nucleic acid including that which encodes, hybridizes under appropriate stringency conditions, or is complementary to a nucleic acid encoding a protein according to the present invention.

As used herein, the terms “hybridization” (hybridizing) and “specificity” (specific for) in the context of nucleotide sequences are used interchangeably. The ability of two nucleotide sequences to hybridize to each other is based upon the degree of complementarity of the two nucleotide sequences, which in turn is based on the fraction of matched complementary nucleotide pairs. The more nucleotides in a given sequence that are complementary to another sequence, the greater the degree of hybridization of one to the other. The degree of hybridization also depends on the conditions of stringency which include temperature, solvent ratios, salt concentrations, and the like. In particular, “selective hybridization” pertains to conditions in which the degree of hybridization of a polynucleotide of the invention to its target would require complete or nearly complete complementarity. The complementarity must be sufficiently high so as to assure that the polynucleotide of

the invention will bind specifically to the target nucleotide sequence relative to the binding of other nucleic acids present in the hybridization medium. With selective hybridization, complementarity will be about 90-100%, preferably about 95-100%, more preferably about 100%.

"Stringent conditions" are those that (1) employ low ionic strength and high temperature for washing, for example: 0.015 M NaCl, 0.0015 M sodium titrate, 0.1% SDS at 50°C; or (2) employ during hybridization a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin, 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42°C. Another example is use of 50% formamide, 5X SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5X Denhardt's solution, sonicated salmon sperm DNA (50 μg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2X SSC and 0.1% SDS. A skilled artisan can readily determine and vary the stringency conditions appropriately to obtain a clear and detectable hybridization signal.

As used herein, a nucleic acid molecule is said to be "isolated" or "purified" when the nucleic acid molecule is substantially separated from contaminant nucleic acid encoding other polypeptides from the source of nucleic acid. Isolated or purified is also meant to include nucleic acids which encode Dkk or fragments thereof which lack surrounding genomic sequences that flank the *Dkk* gene. Isolated or purified is further intended to include nucleic acids which encode Dkk interacting proteins or biologically active fragments thereof which lack surrounding genomic sequences that flank the Dkk interacting protein genes.

The present invention further provides fragments of the encoding nucleic acid molecule. As used herein, a fragment of an encoding nucleic acid molecule refers to a small portion of the entire protein encoding sequence. The size of the fragment will be determined by the intended use. For example, if the fragment is chosen so as to encode an active portion of the protein, the fragment will need to be large enough to encode the functional region(s) of the protein. If the fragment is to be used as a

nucleic acid probe or PCR primer, then the fragment length is chosen so as to obtain a relatively small number of false positives during probing/priming.

Fragments of the encoding nucleic acid molecules of the present invention (*i.e.*, synthetic oligonucleotides) that are used as probes or specific primers for the polymerase chain reaction (PCR), or to synthesize gene sequences encoding proteins of the invention can easily be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci *et al.* (*J. Am. Chem. Soc.* 103:3185-3191 (1981)) or using automated synthesis methods. In addition, larger DNA segments can readily be prepared by well known methods, such as synthesis of a group of oligonucleotides that define various modular segments of the gene, followed by ligation of oligonucleotides to build the complete modified gene.

The polypeptide encoding nucleic acid molecules of the present invention may further be modified to contain a detectable label for diagnostic and probe purposes. A variety of such labels are known in the art and can readily be employed with the encoding molecules herein described. Suitable labels include, but are not limited to, biotin, radiolabeled nucleotides and the like. A skilled artisan can employ any of the art known labels to obtain a labeled encoding nucleic acid molecule.

Modifications to the primary structure itself by deletion, addition, or alteration of the amino acids incorporated into the protein sequence during translation can be made without destroying the activity of the protein. Such substitutions or other alterations result in proteins having an amino acid sequence encoded by a nucleic acid falling within the contemplated scope of the present invention.

Antisense molecules corresponding to the polypeptide coding or complementary sequence may be prepared. Methods of making antisense molecules which bind to mRNA, form triple helices or are enzymatically active and cleave TSG RNA and single stranded DNA (ssDNA) are known in the art. See, *e.g.*, <u>Antisense and Ribozyme Methodology:Laboratory Companion</u> (Ian Gibson, ed., Chapman & Hall, 1997) and <u>Ribozyme Protocols: Methods in Molecular Biology</u> (Phillip C. Turner, ed., Humana Press, Clifton, NJ, 1997).

Also contemplated is the use of compounds which mediate postranscriptional gene silencing (PTGS), quelling and RNA interference (RNAi). These compounds typically are about 21 to about 25 nucleotides and are also known as short interfering RNAs or short inhibitory RNAs (siRNAs). The siRNAs are produced from an initiating double stranded RNA (dsRNA). Although the full mechanism by which the siRNAs function is not fully elucidated, it is known that these siRNAs transform the target mRNA into dsRNA, which is then degraded. Preferred forms are 5' phosphorylated siRNAs, however, hydroxylated forms may also be utilized. For additional background regarding the preparation and mechanism of siRNAs generally, see, e.g., Lipardi *et al., Cell* 107(3): 297-307 (2001); Boutla *et al., Curr. Biol.* 11(22): 1776-80 (2001); Djikeng *et al., RNA* 7(11): 1522-30 (2001); Elbashir *et al., EMBO J.* 20(23): 6877-88 (2001); Harborth *et al., J. Cell. Sci.* 114(Pt. 24): 4557-65 (2001); Hutvagner *et al., Science* 293(5531): 811-3 (2001); and Elbashir *et al., Nature* 411:494-98 (2001).

Also contemplated are short hairpin RNAs (shRNAs). shRNAs are a modification of the siRNA method described above. Instead of transfecting exogenously synthesized dsRNA into a cell, sequence-specific silencing can be achieved by stabling expressing siRNA from a DNA template as a fold-back stem-loop, or hairpin. This approach is known as shRNA. This method permits the analysis of loss of function phenotypes due to sequence-specific gene silencing in mammalian cells by avoiding many of the problems associated with siRNAs, such as RNase degradation of the reagents, expensive chemical synthesis, etc. For additional background regarding the preparation and mechanism of shRNAs generally, see, e.g., Yu *et al., PNAS* 99:6047-6052 (2002); Paddison *et al., Genes and Devel.* 16:948-58 (2002); and Brummelkamp *et al.*, Science 296:550-553 (2002). For additional background on the use of this method in mammalian gene knockdown methodologies, see Tuschl, *Nature Biotech.* 20:446-448 (2002) (and references therein).

In one preferred embodiment, the siRNA or shRNA is directed to a Dkk encoding mRNA, wherein a preferred Dkk is Dkk-1. In another embodiment, the

siRNA or shRNA is directed towards a protein which binds to and modulates the activity of or is modulated by a Dkk; these proteins include LRP5, LRP6 and HBM as well as other members of the Wnt pathway.

6. <u>Isolation of Other Related Nucleic Acid Molecules</u>

The identification of the nucleic acid molecule of Dkk allows a skilled artisan to isolate nucleic acid molecules that encode other members of the Dkk family (*see*, Krupnik *et al.*, 1999). Further, the presently disclosed nucleic acid molecules allow a skilled artisan to isolate nucleic acid molecules that encode Dkk-1-like proteins, in addition to Dkk-1. The presently disclosed Dkk-1 interacting proteins and their corresponding nucleic acid molecules allows a skilled artisan to further isolate other related protein family members which interact with Dkk-1.

A skilled artisan can readily use the amino acid sequence of Dkk and Dkk interacting proteins to generate antibody probes to screen expression libraries prepared from appropriate cells. Typically, polyclonal antiserum from mammals such as rabbits immunized with the purified protein (as described below) or monoclonal antibodies can be used to probe a mammalian cDNA or genomic expression library, such as a human macrophage library, to obtain the appropriate coding sequence for other members of the protein family. The cloned cDNA sequence can be expressed as a fusion protein, expressed directly using its own control sequences, or expressed by constructions using control sequences appropriate to the particular host used for expression of the desired protein.

Alternatively, a portion of the coding sequence herein described can be synthesized and used as a probe to retrieve DNA encoding a member of the protein family from any mammalian organism. Oligomers containing approximately 18-20 nucleotides (encoding about a 6-7 amino acid stretch) are prepared and used to screen genomic DNA or cDNA libraries to obtain hybridization under stringent conditions or conditions of sufficient stringency to eliminate an undue level of false positives.

Additionally, pairs of oligonucleotide primers can be prepared for use in a polymerase chain reaction (PCR) to selectively clone an encoding nucleic acid

molecule. A PCR denature/anneal/extend cycle for using such PCR primers is well known in the art and can readily be adapted for use in isolating other encoding nucleic acid molecules. For example, degenerate primers can be utilized to obtain sequences related to Dkk-1 or Dkk-1 interacting proteins. Primers can be designed that are not perfectly complementary and can still hybridize to a portion of a target sequence or flanking sequence and thereby provide for amplification of all or a portion of a target sequence. Primers of about 20 nucleotides or less, preferably have about one to three mismatches located at the 5' and/or 3' ends. Primers of about 20 to 30 nucleotides have up to about 30% mismatches and can still hybridize to a target sequence. Hybridization conditions for primers with mismatch can be determined by the method described in Maniatis *et al.*, Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982) or by reference to known methods. The ability of the primer to hybridize to a sequence of either Dkk-1, a Dkk-1 interacting protein, or a related sequence under varying conditions can be determined using this method. Because a target sequence is known, the effect of mismatches can be determined by methods known to those of skill in the art. Degenerate primers would be based on putative conserved amino acid sequences of the Dkk-1 and Dkk-1 interacting protein genes.

7. rDNA Molecules for Polypeptide Expression

The present invention further provides recombinant DNA molecules (rDNAs) that contain a polypeptide coding sequence. As used herein, a rDNA molecule is a DNA molecule that has been subjected to molecular manipulation *in situ*. Methods for generating rDNA molecules are well known in the art, for example, see Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1989). In the preferred rDNA molecules, a coding DNA sequence is operably linked to expression control sequences and/or vector sequences.

The choice of vector and/or expression control sequences to which one of the protein family encoding sequences of the present invention is operably linked depends directly, as is well known in the art, on the functional properties desired, *e.g.*, protein

expression, and the host cell to be transformed. A vector contemplated by the present invention is at least capable of directing the replication and/or insertion into the host chromosome, and preferably also expression, of the structural gene included in the rDNA molecule.

Expression control elements that are used for regulating the expression of an operably linked protein encoding sequence are known in the art and include, but are not limited to, inducible promoters, constitutive promoters, secretion signals, and other regulatory elements. Preferably, the inducible promoter is readily controlled, such as being responsive to a nutrient in the host cell's medium. Preferred promoters include yeast promoters, which include promoter regions for metallothionein, 3-phosphoglycerate kinase or other glycolytic enzymes such as enolase or glyceraldehyde-3-phosphate dehydrogenase, enzymes responsible for maltose and galactose utilization, and others. Vectors and promoters suitable for use in yeast expression are further described in EP 73,675A. Appropriate non-native mammalian promoters might include the early and late promoters from SV40 (Fiers et al, *Nature*, 273:113 (1978)) or promoters derived from Moloney murine leukemia virus, mouse tumor virus, avian sarcoma viruses, adenovirus II, bovine papilloma virus or polyoma. In addition, the construct may be joined to an amplifiable gene (*e.g.*, DHFR) so that multiple copies of the gene may be made. For appropriate enhancer and other expression control sequences, see also Enhancers and Eukaryotic Gene Expression (Cold Spring Harbor Press, Cold Spring Harbor, NY, 1983). Preferred bone related promoters include CMVbActin or type I collagen promoters to drive expression of the human HBM, Zmax1/LRP5 or LRP6 cDNA. Other preferred promoters for mammalian expression are from cytomegalovirus (CMV), Rous sarcoma virus (RSV), Simian virus 40 (SV40), and EF-1a (human elongation factor 1a-subunit).

In one embodiment, the vector containing a coding nucleic acid molecule will include a prokaryotic replicon, *i.e.*, a DNA sequence having the ability to direct autonomous replication and maintenance of the recombinant DNA molecule extrachromosomally in a prokaryotic host cell, such as a bacterial host cell, transformed therewith. Such replicons are well known in the art. In addition, vectors with a

prokaryotic replicon may also include a gene whose expression confers a detectable marker such as a drug resistance. Typical bacterial drug resistance genes are those that confer resistance to ampicillin or tetracycline.

Vectors that include a prokaryotic replicon can further include a prokaryotic or bacteriophage promoter capable of directing the expression (transcription and translation) of the coding gene sequences in a bacterial host cell, such as *E. coli*. A promoter is an expression control element formed by a DNA sequence that permits binding of RNA polymerase and transcription to occur. Promoter sequences compatible with bacterial hosts are typically provided in plasmid vectors containing convenient restriction sites for insertion of a DNA segment of the present invention. Typical of such vector plasmids are pUC8, pUC9, pBR322 and pBR329 available from Biorad Laboratories, (Richmond, CA), and pPL and pKK223 available from Pharmacia (Piscataway, NJ).

Expression vectors compatible with eukaryotic cells, preferably those compatible with vertebrate cells, can also be used to form a rDNA molecule that contains a coding sequence. Eukaryotic cell expression vectors are well known in the art and are available from several commercial sources. Typically, such vectors are provided containing convenient restriction sites for insertion of a desired DNA segment. Typical of such vectors are pSVL and pKSV-10 (Pharmacia), pBPV-1/pML2d (International Biotechnologies, Inc.), vector systems that include Histidine Tags and periplasmic secretion, or other vectors described in the art.

Eukaryotic cell expression vectors used to construct the rDNA molecules of the present invention may further include a selectable marker that is effective in an eukaryotic cell, preferably a drug resistance selection marker. A preferred drug resistance marker is the gene whose expression results in neomycin resistance, *i.e.*, the neomycin phosphotransferase (*neo*) gene (Southern *et al.*, *J. Mol. Anal. Genet.* 1:327-341 (1982)). Alternatively, the selectable marker can be present on a separate plasmid, and the two vectors introduced by co-transfection of the host cell, and selected by culturing in the appropriate drug for the selectable marker.

8. Host Cells Containing an Exogenously Supplied rDNA Nucleic Acid Molecule

The present invention further provides host cells transformed with a nucleic acid molecule that encodes a polypeptide or protein of the present invention. The host cell can be either prokaryotic or eukaryotic. Eukaryotic cells useful for expression of a protein of the invention are not limited, so long as the cell line is compatible with cell culture methods and compatible with the propagation of the expression vector and expression of the gene product. Preferred eukaryotic host cells include, but are not limited to, yeast, insect and mammalian cells, preferably vertebrate cells such as those from a mouse, rat, monkey or human cell line but also can include invertebrates with, for example, cartilage. Preferred eukaryotic host cells include but are not limited to Chinese hamster ovary (CHO) cells (ATCC No. CCL61), NIH Swiss mouse embryo cells NIH/3T3 (ATCC No. CRL 1658), baby hamster kidney cells (BHK), HOB-03-CE6 osteoblast cells, and other like eukaryotic tissue culture cell lines.

Any prokaryotic host can be used to express a rDNA molecule encoding a protein of the invention. A preferred prokaryotic host is *E. coli*.

Transformation of appropriate cell hosts with a recombinant DNA (rDNA) molecule of the present invention is accomplished by well known methods that typically depend on the type of vector used and host system employed. With regard to transformation of prokaryotic host cells, electroporation and salt treatment methods are typically employed; see, for example, Cohen *et al.*, *Proc. Natl. Acad. Sci. USA* 69: 2110 (1972); Maniatis *et al.* (1982); and Sambrook *et al.* (1989). With regard to transformation of vertebrate cells with vectors containing rDNAs, electroporation, cationic lipid or salt treatment methods are typically employed; see, for example, Graham *et al.*, *Virol.* 52: 456 (1973); Wigler *et al.*, *Proc. Natl. Acad. Sci. USA* 76: 1373-76 (1979).

Successfully transformed cells, *i.e.*, cells that contain a rDNA molecule of the present invention, can be identified by well known techniques including the selection for a selectable marker. For example, cells resulting from the introduction of an rDNA of the present invention can be cloned to produce single colonies. Cells from those

colonies can be harvested, lysed and their DNA content examined for the presence of the rDNA using a method such as that described by Southern, *J. Mol. Biol.* 98: 503 (1975), or Berent *et al.*, *Biotech.* 3: 208 (1985). Alternatively, the cells can be cultured to produce the proteins encoded by the rDNA and the proteins harvested and assayed, using for example, any suitable immunological method. See, *e.g.*, Harlow *et al.*, (1988).

Recombinant DNA can also be utilized to analyze the function of coding and non-coding sequences. Sequences that modulate the translation of the mRNA can be utilized in an affinity matrix system to purify proteins obtained from cell lysates that associate with the Dkk-1 or Dkk-1 interacting protein or expression control sequence. Synthetic oligonucleotides would be coupled to the beads and probed with the lysates, as is commonly known in the art. Associated proteins could then be separated using, for example, a two dimensional SDS-PAGE system. Proteins thus isolated could be further identified using mass spectroscopy or protein sequencing. Additional methods would be apparent to the skilled artisan.

9. <u>Production of Recombinant Peptides and Proteins using a cDNA or Other Recombinant Nucleic Acids</u>

The invention also relates to nucleic acid molecules which encode a Dkk protein and polypeptide fragments thereof, and proteins and polypeptides which bind to Dkk (*e.g.*, LRP5, LRP6 and HBM, Dkk interacting proteins such as the proteins of Figure 5) and molecular analogues. The polypeptides of the present invention include the full length Dkk and polypeptide fragments thereof, Dkk binding proteins and polypeptides thereof. Preferably these proteins are mammalian proteins, and most preferably human proteins and biologically active fragments thereof. Alternative embodiments include nucleic acid molecules encoding polypeptide fragments having a consecutive amino acid sequence of at least about 3, 5, 7, 8, 9, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 175, or 200 amino acid residues from a common polypeptide sequence; amino acid sequence variants of a common polypeptide sequence wherein an amino acid residue has been inserted N- or C-terminal to, or within, the polypeptide sequence or its fragments; and amino acid sequence variants of the common

polypeptide sequence or its fragments, which have been substituted by another conserved residue. Recombinant nucleic acid molecules which encode polypeptides include those containing predetermined mutations by, *e.g.*, homologous recombination, site-directed or PCR mutagenesis, and recombinant Dkk proteins or polypeptide fragments of other animal species, including but not limited to vertebrates (*e.g.*, rabbit, rat, murine, porcine, camelid, reptilian, caprine, avian, fish, bovine, ovine, equine and non-human primate species) as well as invertebrates, and alleles or other naturally occurring variants and homologs of Dkk binding proteins of the foregoing species and of human sequences. Also contemplated herein are derivatives of the commonly known Dkk, Dkk interacting proteins, or fragments thereof, wherein Dkk, Dkk interacting proteins, or their fragments have been covalently modified by substitution, chemical, enzymatic, or other appropriate means with a moiety other than a naturally occurring amino acid (for example a detectable moiety such as an enzyme or radioisotope) and soluble forms of Dkk. It is further contemplated that the present invention also includes nucleic acids with silent mutations which will hybridize to the endogenous sequence and which will still encode the same polypeptide.

The nucleic acid molecules encoding Dkk binding proteins, the LRP5 binding domain fragment of Dkk, or other polypeptides of the present invention are preferably those which share a common biological activity (*e.g.*, mediate Dkk activity such as its interaction with LRP5, HBM or LRP6). The polypeptides of the present invention include those encoded by a nucleic acid molecule with silent mutations, as well as those nucleic acids encoding a biologically active protein with conservative amino acid substitutions, allelic variants, and other variants of the disclosed polypeptides which maintain at least one Dkk activity.

The amino acid compounds of the invention are polypeptides which are partially defined in terms of amino acid residues of designated classes. Polypeptide homologs would include conservative amino acid substitutions within the amino acid classes described below. Amino acid residues can be generally sub-classified into four major subclasses as follows:

Acidic: The residue has a negative charge due to loss of H^+ ion at physiological pH, and the residue is attracted by aqueous solution so as to seek the surface positions in the conformation of a peptide in which it is contained when the peptide is in aqueous medium, at physiological pH.

Basic: The residue has a positive charge due to association with H^+ ion at physiological pH, and the residue is attracted by aqueous solution so as to seek the surface positions in the conformation of a peptide in which it is contained when the peptide is in aqueous medium at physiological pH.

Neutral/non-polar: The residues are not charged at physiological pH, but the residue is repelled by aqueous solution so as to seek the inner positions in the conformation of a peptide in which it is contained when the peptide is in aqueous medium. These residues are also designated "hydrophobic."

Neutral/polar: The residues are not charged at physiological pH, but the residue is attracted by aqueous solution so as to seek the outer positions in the conformation of a peptide in which it is contained when the peptide is in aqueous medium.

It is understood, of course, that in a statistical collection of individual residue molecules some molecules will be charged, and some not, and there will be an attraction for or repulsion from an aqueous medium to a greater or lesser extent. To fit the definition of "charged", a significant percentage (at least approximately 25%) of the individual molecules are charged at physiological pH. The degree of attraction or repulsion required for classification as polar or nonpolar is arbitrary and, therefore, amino acids specifically contemplated by the invention have been classified as one or the other. Most amino acids not specifically named can be classified on the basis of known behavior.

Amino acid residues can be further subclassified as cyclic or noncyclic, and aromatic or non-aromatic, self-explanatory classifications with respect to the side chain substituent groups of the residues, and as small or large. The residue is considered small if it contains a total of 4 carbon atoms or less, inclusive of the carboxyl carbon. Small residues are, of course, always nonaromatic.

The gene-encoded secondary amino acid proline, although technically within the group neutral/nonpolar/large/cyclic and nonaromatic, is a special case due to its known effects on the secondary conformation of peptide chains, and is not, therefore, included in this defined group.

Other amino acid substitutions of those encoded in the gene can also be included in peptide compounds within the scope of the invention and can be classified within this general scheme according to their structure.

All of the compounds of the invention may be in the form of the pharmaceutically acceptable salts or esters. Salts may be, for example, Na^+, K^+, Ca^{+2}, Mg^{+2} and the like; the esters are generally those of alcohols of 1-6 carbons.

The present invention further provides methods for producing a protein of the invention using nucleic acid molecules herein described. In general terms, the production of a recombinant form of a protein typically involves the following steps.

First, a nucleic acid molecule is obtained that encodes Dkk, such as a nucleic acid molecule encoding human Dkk or any other Dkk sequence, or that encodes a Dkk binding protein, a Dkk aptamer or a biologically active fragment thereof. Particularly for Dkk binding peptides, the nucleotides encoding the peptide are incorporated into a nucleic acid in the form of an in-frame fusion, insertion into or appended to a thioredoxin coding sequence. The coding sequence (ORF) is directly suitable for expression in any host, as it is not interrupted by introns.

These DNAs can be transfected into host cells such as eukaryotic cells or prokaryotic cells. Eukaryotic hosts include mammalian cells and vertebrate (*e.g.*, osteoblasts, osteosarcoma cell lines, Drosophila S2 cells, hepatocytes, tumor cell lines and other bone cells of any mammal, as well as insect cells, such as Sf9 cells using recombinant baculovirus). For example, a DNA expressing an open reading frame (ORF) under control of a type I collagen promoter, or such osteoblast promoters as osteocalcin histone, type I collagen, TGFβ1, MSX2, cfos/cJun and Cbfa1, can be used to regulate the Dkk in animal cells. Alternatively, the nucleic acid can be placed downstream from an inducible promoter, which can then be placed into vertebrate or invertebrate cells or be used in creating a transgenic animal model.

Alternatively, proteins and polypeptides of the present invention can be expressed in an heterologous system. The human cell line GM637, SV-40 transformed human fibroblasts, can be transfected, with a plasmid containing a Dkk ligand binding domain coding sequence under the control of the chicken actin promoter (Reis *et al.*, EMBO J. 11: 185-193 (1992)). Such transfected cells could be used as a source of Dkk binding domain in functional assays. Alternatively, polypeptides encoding only a portion of Dkk or any of the disclosed Dkk binding peptides Dkk aptamers or a polypeptide encoding a Dkk interacting protein can be expressed alone or in the form of a fusion protein. For example, Dkk derived peptides can be expressed in bacteria (*e.g.*, *E. coli*) as GST- or His-Tag fusion proteins. These fusion proteins are then purified and can be used to generate polyclonal antibodies or can be used to identify other Dkk ligands.

The nucleic acid coding sequence is preferably placed in operable linkage with suitable control sequences, as described above, to form an expression unit containing the protein encoding open reading frame. The expression unit is used to transform a suitable host and the transformed host is cultured under conditions that allow the production of the recombinant protein. Optionally the recombinant protein is isolated from the medium or from the cells; recovery and purification of the protein may not be necessary in some instances where some impurities may be tolerated.

Each of the foregoing steps can be done in a variety of ways. For example, the desired coding sequences may be obtained from genomic fragments and used directly in appropriate hosts. The construction of expression vectors that are operable in a variety of hosts is accomplished using appropriate replicons and control sequences, as set forth above. The control sequences, expression vectors, and transformation methods are dependent on the type of host cell used to express the gene and were discussed in detail earlier. Suitable restriction sites can, if not normally available, be added to the ends of the coding sequence so as to provide an excisable gene to insert into these vectors. A skilled artisan can readily adapt any host/expression system known in the art for use with the nucleic acid molecules of the invention to produce recombinant protein.

10. Methods to Identify Binding Partners

Another embodiment of the present invention provides methods for use in isolating and identifying binding partners of Dkk or Dkk interacting proteins. Dkk or a Dkk interacting protein or a polypeptide fragment thereof can be mixed with a potential binding partner or an extract or fraction of a cell under conditions that allow the association of potential binding partners with Dkk or with Dkk interacting proteins. After mixing, the peptides, polypeptides, proteins or other molecules that have become associated with Dkk or a Dkk interacting protein are separated from the mixture. The binding partner that bound to the polypeptide then can be purified and further analyzed. Determination of binding partners of Dkk and Dkk interacting proteins as well as agents which prevent the interaction of Dkk with one of its interacting proteins (*e.g.*, LRP5, LRP6, HBM, or those proteins listed in Figure 5) can be performed using a variety of different competition assays as are known in the art. For example, the minimal sequence of Dkk, as described herein, can be used to identify antibodies which compete with LRP5 (or LRP6, HBM or other ligand binding partners) for binding to Dkk-1 and vice versa. The minimal Dkk sequence can be bound to the bottom of a 96-well plate (or other solid substrate), and antibodies or other potential binding agents (*e.g.*, polypeptides, mimetics, homologs, antibody fragments and the like) can be screened in a competition assay to identify agents with binding affinities, for example, greater than the natural ligand binding partner of Dkk.

In the present invention, suitable cells are used for preparing assays, for the expression of a LRP and/or Dkk or proteins that interact therewith. The cells may be made or derived from mammals, yeast, fungi, or viruses. A suitable cell for the purposes of this invention is one that includes but is not limited to a cell that can exhibit a detectable Dkk-LRP (or HBM) interaction, and preferably, the differential interaction between Dkk-1-LRP5 and Dkk-1-HBM. For the desired assay, the cell type may vary. In several embodiments, bone cells are preferred, for example, a human osteoblast cell (e.g. hOB-03-CE6) or osteosarcoma cell (e.g. U2OS). Additional hOB cells are hOB-03-C5, hOB-02-02 and, an immortalized pre-osteocytic cell line referred to as hOB-01-C1-PS-09 cells (which are deposited with American Type Culture Collection in

Manassas, Va. with the designation PTA-785), Examples of osteosarcoma cells would include SaoS2, MG63 and HOS TE85 Immortalized refers to a substantially continuous and permanently established cell culture with substantially unlimited cell division potential. That is, the cells can be cultured substantially indefinitely, i.e., for at least about 6 months under rapid conditions of growth, preferably much longer under slower growth conditions, and can be propagated rapidly and continually using routine cell culture techniques. Alternatively stated, preferred cells can be cultured for at least about 100, 150 or 200 population doublings. These cells produce a complement of proteins characteristic of normal human osteoblastic cells and are capable of osteoblastic differentiation. They can be used in cell culture studies of osteoblastic cell sensitivity to various agents, such as hormones, cytokines, and growth factors, or in tissue therapy. Certain non bone cells such as HEK 293 cells that exhibit detectable Dkk-LRP (or HBM) interaction are also be useful for the assays of this invention.

To identify and isolate a binding partner, the entire Dkk protein (*e.g.*, human Dkk-1, GenBank Accession No. BAA34651) or a Dkk interacting protein (Genbank Accession Nos. for some Dkk-1 interacting proteins are given in Figure 5) can be used. Alternatively, a polypeptide fragment of the protein can be used. Suitable fragments of the protein include at least about 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150 or more contiguous amino acid residues of any Dkk or Dkk interactor sequence. Preferable sequences of Dkk include portions or all of one or both of the cysteine rich domains (*e.g.*, Cys-1 and Cys-2 of Dkk-1) or the conserved sequences at the amino terminus of Dkk-1 (*See* Krupnik *et al.*, *Gene* 238: 301-313 (1999)). Alternatively, portions of LRP5, LRP6, HBM and other Dkk interacting proteins such as those in Figure 5 that interact with Dkk-1 can be used to identify and isolate agents which modulate Dkk activity. Alternatively, peptide aptamers of LRP5, LRP6, HBM, Dkk and other Dkk interacting proteins such as those in Figure 5 that interact with Dkk-1 can be used to identify and isolate agents which modulate Dkk activity.

As used herein, a cellular extract refers to a preparation or fraction which is made from a lysed or disrupted cell. A variety of methods can be used to obtain cell

extracts. Cells can be disrupted using either physical or chemical disruption methods. Examples of physical disruption methods include, but are not limited to, sonication and mechanical shearing. Examples of chemical lysis methods include, but are not limited to, detergent lysis and enzyme lysis. A skilled artisan can readily adapt methods for preparing cellular extracts in order to obtain extracts for use in the present methods.

Once an extract of a cell is prepared, the extract is mixed with the protein of the invention under conditions in which association of the protein with the binding partner can occur. A variety of conditions can be used, the most preferred being conditions that closely resemble conditions found in the cytoplasm of a human cell. Features such as osmolarity, pH, temperature, and the concentration of cellular extract used, can be varied to optimize the association of the protein with the binding partner.

After mixing under appropriate conditions, the bound complex is separated from the mixture. A variety of techniques can be utilized to separate the mixture. For example, antibodies specific to a protein of the invention can be used to immunoprecipitate the binding partner complex. Alternatively, standard chemical separation techniques such as chromatography and density/sediment centrifugation can be used. For example, a protein of the invention is expressed with an affinity tag such as a His tag. The His labeled protein and any bound molecule may be retained and selectively eluted from a Ni-NTA column.

After removal of non-associated cellular constituents found in the extract, the binding partner can be dissociated from the complex using conventional methods. For example, dissociation can be accomplished by altering the salt concentration or pH of the mixture.

To aid in separating associated binding partner pairs from the mixed extract, the protein of the invention can be immobilized on a solid support. For example, the protein can be attached to a nitrocellulose matrix or acrylic beads. Attachment of the protein to a solid support aids in separating peptide/binding partner pairs from other constituents found in the extract. The identified binding partners can be either a single protein or a complex made up of two or more proteins.

Alternatively, the nucleic acid molecules of the invention can be used in a Y2H system. The Y2H system has been used to identify other protein partner pairs and can readily be adapted to employ the nucleic acid molecules herein described. Methods of performing and using Y2H systems are known. See, *e.g.*, Finley *et al.*, "Two-Hybrid Analysis of Genetic Regulatory Networks," in The Yeast Two-Hybrid System (Paul L. Bartel et al., eds., Oxford, 1997); Meijia Yang, "Use of a Combinatorial Peptide Library in the Two-Hybrid Assay," in The Yeast Two-Hybrid System (Paul L. Bartel et al., eds., Oxford, 1997); Gietz *et al.*, "Identification of proteins that interact with a protein of interest: Applications of the yeast two-hybrid system," *Mol. & Cell. Biochem.* 172: 67-9 (1997); K. H. Young, "Yeast Two-Hybrid: So Many Interactions,(in) so Little Time," *Biol. Reprod.* 58: 302-311 (1998); R. Brent *et al.*, "Understanding Gene and Allele Function with Two-Hybrid Methods," *Annu. Rev. Genet.* 31:663-704 (1997) and U.S. Patent No. 5,989,808. The Dkk-1 interacting proteins identified in Figure 5 were identified using the Y2H interacting system using Dkk-1 as bait.

One preferred *in vitro* binding assay for Dkk modulators would comprise a mixture of a LRP binding domain of Dkk and one or more candidate binding targets or substrates. After incubating the mixture under appropriate conditions, one would determine whether Dkk or a fragment thereof bound with the candidate modulator present. For cell-free binding assays, one or more of the components usually comprises or is coupled to a label. The label may provide for direct detection, such as radioactivity, luminescence, optical or electron density, *etc.*, or indirect detection such as an epitope tag, an enzyme, *etc.* A variety of methods may be employed to detect the label depending on the nature of the label and other assay components. For example, the label may be detected bound to the solid substrate or a portion of the bound complex containing the label may be separated from the solid substrate, and the label thereafter detected. Fluorescence resonance energy transfer may be utilized to monitor the interaction of two labeled molecules. For example, a fluorescence label on Dkk and another label on LRP5 or a soluble fragment thereof such as the extracellular domain will exchange fluorescence resonance energy when in close proximity indicating that the two molecules are bound. A preferred binding partner for Dkk will

increase or decrease the affinity between Dkk and LRP5 which will be readily observable in a fluorescence spectrometer. Alternatively, an instrument, such as a surface plasmon resonance detector manufactured by BIAcore (Uppsala, Sweden), may be used to observe interactions with a fixed target. One skilled in the art knows of many other methods which may be employed for this purpose.

Thereby, the present invention provides methods for screening candidates including polypeptides of the present invention for activity which identifies these candidates as valuable drug leads. Other suitable methods are also known in the art and are suitable for use herein, including *Xenopus* oocyte injection studies and TCF luciferase assays.

Additional assays can be used to identify the activity of Dkk and Dkk interacting proteins in the Wnt pathway, as well as the impact of modulators of Dkk and Dkk interacting proteins on the Wnt pathway. These include, for example, a *Xenopus* embryo assay and a TCF-luciferase reporter gene assay to monitor Wnt signaling modulation.

Xenopus embryos are an informative *in vivo* assay system to evaluate the modulation of Wnt signaling. Ectopic expression of certain Wnts or other activators of the Wnt signaling pathway results in a bifurcation of the anterior neural plate. This bifurcation results in a duplicated body axis, which suggests a role for Wnt signaling during embryonic development (McMahon *et al.*, *Cell* 58: 1075-84 (1989); Sokol *et al.*, *Cell* 67: 741-52 (1991)). Since these original observations, the *Xenopus* embryo assay has been extensively used as an assay system for evaluating modulation of the Wnt signaling pathway. One preferred embodiment of the present invention is demonstrated in Example 6.

Constructs for *Xenopus* expression can be prepared as would be known in the art. For example, a variety of cDNAs have been engineered into the vector pCS2+ (Turner *et al.*, *Genes Devel.* 8: 1434-1447 (1994)) to facilitate the *in vitro* generation of mRNA for use in *Xenopus* embryo injection experiments. DNA inserts are subcloned in the sense orientation with respect to the vector SP6 promoter. Downstream of the insert, the vector provides an SV40 virus polyadenlylation signal and a T3 promoter

sequence (*i.e.*, for the generation of antisense mRNA). Constructs can be generated for various Dkk family members, LRP5, LRP6, HBM, Dkk-1 interactors, etc. Constructs could be generated in pCS2$^+$ that contain the nucleic acid sequence encoding for the peptide aptamers that were identified in yeast screens. These sequences would be fused to a 5' synthetic translation initiation sequence followed by a canonical signal sequence to ensure that the peptide aptamer would be translated and secreted from the cell.

Once these constructs are made then mRNA can be synthesized and injected into *Xenopus* oocytes. mRNA for microinjection into *Xenopus* embryos is generated by *in vitro* transcription using the cDNA constructs in the pCS2$^+$ vector described above as template. Various amounts of RNA can be injected into the ventral blastomere of the 4- or 8-cell *Xenopus* embryo substantially as described in Moon *et al.*, *Technique-J. of Methods in Cell and Mol. Biol.* 1: 76-89 (1989), and Peng, *Meth. Cell. Biol.* 36: 657-62 (1991).

Previous data has shown that expression of LRP5, in the presence of Wnt5a, results in a Wnt-induced duplicated axis formation in *Xenopus* embryos (Tamai *et al.*, *Nature* 407: 530-535 (2000)). The roles of Dkk-1 and Dkk-2, and Dkk-1 interacting proteins, in modulating the LRP5-mediated Wnt response *in vivo* can be analyzed using, for example, the *Xenopus* embryo. In addition, the peptide aptamers, Dkk interacting proteins, or combinations of the above can be evaluated in a similar manner.

Experiments can also be conducted wherein RNA is injected into the dorsal blastomere to ensure the specificity of the observed phenotypes. Lineage tracing experiments can be performed where a marker gene such as green fluorescent protein (GFP) or LacZ is co-injected with the experimental RNAs. Detecting marker gene expression would identify the targeted cells of the microinjection and aid in elucidating the mechanism of action. In addition to the Wnt signaling components listed above, the point at which HBM acts upon the Wnt pathway can also be analyzed. This can be done by co-injections of various dominant-negative constructs. For example, a dominant negative TCF-3 construct would be useful to demonstrate that the observed axis duplication (and Wnt activation) is mediated via the β-catenin-TCF response. If so,

such a construct would be expected to abolish the observed duplicated axis phenotype. Another example would include a dominant negative Dsh construct. Since Dsh is far upstream in the Wnt signaling pathway, a dominant negative construct should abolish the activation of the Wnt response and the observed axis duplication. If it does not, this would suggest that axis duplication is being induced via a different signaling pathway.

The marker genes of the injected *Xenopus* embryos can be analyzed as follows. Representative embryos are collected at stage 10.5 (11 hours post fertilization) for marker gene analysis. RNA is extracted and purified from the embryos following standard protocols (Sambrook *et al.*, 1989 at 7.16). Marker genes could include the following: Siamois (*i.e.*, Wnt responsive gene), Xnr3 (*i.e.*, Wnt responsive gene), slug (*i.e.*, neural crest marker), Xbra (*i.e.*, early mesoderm marker), HNK-1 (*i.e.*, ectodermal/neural marker), endodermin (*i.e.*, endoderm), Xlhbox8 (*i.e.*, pancreatic), BMP2 and BMP4 (*i.e.*, early mesoderm), XLRP6 (*i.e.*, maternal and zygotic expression, it is also the LRP6 homolog in the frog), EF-1 (*i.e.*, control) and ODC (*i.e.*, control). Induction of marker genes is analyzed and quantitated by RT-PCR/TaqMan®.

This type of marker analysis is excellent to monitor changes in gene expression that result very early in the embryo as a direct result of signaling perturbation. Other experiments could be designed that would monitor changes in gene expression in a more tissue or spatially-restricted fashion. Examples would include the generation of a transgenic *Xenopus* model. For example, Zmax/LRP5 and HBM expression could be under the control of the brachyury or cardiac-actin promoters directing gene expression transiently in the mesoderm or in the somites, respectively. Phenotype analyses of these transgenic *Xenopus* animals would include marker gene analysis/transcriptional profiling (from a restricted tissue source) and histologic examination of the tissue.

A TCF-luciferase assay system such as that described in Example 7 can also be used to monitor Wnt signaling activity, Dkk activity and Dkk interacting protein activity. Constructs for the TCF-luciferase assays can be prepared as would be known in the art. For example, Dkk and Dkk interacting protein peptides, LRP5/LRP6, among others, can be expressed in pcDNA3.1, using Kozak and signal sequences to target peptides for secretion.

Once constructs have been prepared, cells such as osteoblasts and HEK293 cells are seeded in well plates and transfected with construct DNA, CMV beta-galactosidase plasmid DNA, and TCF-luciferase reporter DNA. The cells are then lysed and assayed for beta-galactosidase and luciferase activity to determine whether Dkk, Dkk interacting proteins, or other molecules such as antibodies affect Wnt signaling.

Additional assays for monitoring Wnt signaling activity, Dkk activity, and Dkk interacting protein activity include:

> Modulation of another Wnt-responsive transcription factor, LEF, as visualized by a reporter gene activity. One example includes the activation of the LEF1 promoter region fused to the luciferase reporter gene (Hsu *et al.*, *Mol. Cell. Biol.* 18: 4807-18 (1999)).
>
> Alterations in cell proliferation, cell cycle or apoptosis. There are numerous examples describing Wnt-mediated cellular transformations including Shimizu *et al.*, *Cell. Growth Differ.* 8: 1349-58 (1997).
>
> Stabilization and cellular localization of de-phosphorylated β-catenin as an indicator of Wnt activation (Shimizu et al., 1997).

Additional methods of assaying Wnt signaling, through either the canonical or non-canonical pathways, would be apparent to the artisan of ordinary skill.

11. <u>Methods to Identify Agents that Modulate the Expression of a Nucleic Acid Encoding the Dkk and/or LRP5 Proteins and/or Dkk interacting proteins</u>

Another embodiment of the present invention provides methods for identifying agents that modulate the expression of a nucleic acid encoding Dkk. Such assays may utilize any available means of monitoring for changes in the expression level of the nucleic acids of the invention. As used herein, an agent is said to modulate the expression of Dkk, if it is capable of up- or down-regulating expression of the nucleic acid in a cell (*e.g.*, mRNA).

In one assay format, cell lines that contain reporter gene fusions between the nucleic acid encoding Dkk (or proteins which modulate the activity of Dkk) and any

assayable fusion partner may be prepared. Numerous assayable fusion partners are known and readily available, including but not limited to the firefly luciferase gene and the gene encoding chloramphenicol acetyltransferase (Alam *et al.*, *Anal. Biochem.* 188: 245-254 (1990)). Cell lines containing the reporter gene fusions are then exposed to the agent to be tested under appropriate conditions and time. Differential expression of the reporter gene between samples exposed to the agent and control samples identifies agents which modulate the expression of a nucleic acid encoding Dkk or other protein which modulates Dkk activity. Such assays can similarly be used to determine whether LRP5 and even LRP6 activity is modulated by regulating Dkk activity.

Additional assay formats may be used to monitor the ability of the agent(s) to modulate the expression of a nucleic acid encoding Dkk, alone or Dkk and LRP5, and/or Dkk interacting proteins such as those identified in Figure 5. For instance, mRNA expression may be monitored directly by hybridization to the nucleic acids of the invention. Cell lines are exposed to the agent to be tested under appropriate conditions and time and total RNA or mRNA is isolated by standard procedures such those disclosed in Sambrook *et al.* (1989); Ausubel *et al.*, <u>Current Protocols in Molecular Biology</u> (Greene Publishing Co., NY, 1995); Maniatis *et al.*, <u>Molecular Cloning: A Laboratory Manual</u> (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982); and <u>Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology</u> (Frederick M. Ausubel *et al.*, April 1999).

Probes to detect differences in RNA expression levels between cells exposed to the agent and control cells may be prepared from the nucleic acids of the invention. It is preferable, but not necessary, to design probes which hybridize only with target nucleic acids under conditions of high stringency. Only highly complementary nucleic acid hybrids form under conditions of high stringency. Accordingly, the stringency of the assay conditions determines the amount of complementarity which should exist between two nucleic acid strands in order to form a hybrid. Stringency should be chosen to maximize the difference in stability between the probe:target hybrid and potential probe:non-target hybrids.

Probes may be designed from the nucleic acids of the invention through methods known in the art. For instance, the G+C content of the probe and the probe length can affect probe binding to its target sequence. Methods to optimize probe specificity are commonly available. See for example, Sambrook *et al.* (1989) or Ausubel *et al.* (Current Protocols in Molecular Biology, Greene Publishing Co., NY, 1995).

Hybridization conditions are modified using known methods, such as those described by Sambrook *et al.* (1989) and Ausubel *et al.* (1995), as suitable for each probe. Hybridization of total cellular RNA or RNA enriched for polyA RNA can be accomplished in any available format. For instance, total cellular RNA or RNA enriched for polyA RNA can be affixed to a solid support and the solid support exposed to at least one probe comprising at least one, or part of one of the nucleic acid sequences of the invention under conditions in which the probe will specifically hybridize. Alternatively, nucleic acid fragments comprising at least one, or part of one of the sequences of the invention can be affixed to a solid support, such as a porous glass wafer. The glass or silica wafer can then be exposed to total cellular RNA or polyA RNA from a sample under conditions in which the affixed sequences will specifically hybridize. Such glass wafers and hybridization methods are widely available, for example, those disclosed by Beattie (WO 95/11755). By examining for the ability of a given probe to specifically hybridize to an RNA sample from an untreated cell population and from a cell population exposed to the agent, agents which up- or down-regulate the expression of a nucleic acid encoding Dkk, a Dkk interacting protein, and/or LRP5 can be identified.

Microarray technology and transcriptional profiling are examples of methods which can be used to analyze the impact of putative Dkk or Dkk interacting protein modulating compounds. For transcriptional profiling, mRNA from cells exposed *in vivo* to a potential Dkk modulating agent, such as the Dkk interacting proteins identified in the present invention (*e.g.*, those identified in Figure 5), agents which modulate Dkk interacting proteins, and mRNA from the same type of cells that were not exposed to the agent could be reverse transcribed and hybridized to a chip containing DNA from

numerous genes, to thereby compare the expression of genes in cells treated and not treated with the agent. If, for example a putative Dkk modulating agent down-regulates the expression of Dkk in the cells, then use of the agent may be undesirable in certain patient populations. For additional methods of transcriptional profiling and the use of microarrays, refer to, for example, U.S. Patent No. 6,124,120 issued to Lizardi (2000).

Additional methods for screening the impact of Dkk and Dkk interacting protein modulating compounds or the impact of Dkk or Dkk interacting proteins on modulation of LRP5, LRP6, HBM or the Wnt pathway include the use of TaqMan PCR, conventional reverse transcriptase PCR (RT-PCR), changes in downstream surrogate markers (*i.e.*, Wnt responsive genes), and anti-Dkk Western blots for protein detection. Other methods would be readily apparent to the artisan of ordinary skill.

12. Methods to Identify Agents that Modulate at Least One Activity of Dkk, a Dkk Interacting Protein, or LRP5/LRP6/HBM

Another embodiment of the present invention provides methods for identifying agents that modulate at least one activity of Dkk, Dkk interacting proteins, and/or LRP5/LRP6/HBM proteins or preferably which specifically modulate an activity of a Dkk/Dkk interacting protein complex or an LRP5(or LRP6/HBM)/Dkk complex, or a biologically active fragment of Dkk (*e.g.*, comprising the domain which binds LRP5/LRP6/HBM) or a Dkk interacting protein complex. Such methods or assays may utilize any means of monitoring or detecting the desired activity as would be known in the art (*See, e.g.*, Wu *et al.*, *Curr. Biol.* 10:1611-4 (2000); Fedi *et al.*, *J. Biol. Chem.* 274:19465-72 (1991); Grotewold *et al.*, *Mech. Dev.* 89:151-3 (1999); Shibata *et al.*, *Mech. Dev.* 96:243-6 (2000); Wang *et al.*, *Oncogene* 19:1843-8 (2000); and Glinka *et al.*, *Nature* 391:357-62 (1998)). Potential agents which modulate Dkk include, for example, p53, the tumor suppressor protein, which can induce Dkk-1. Damage to DNA has also been observed to up-regulate Dkk-1 expression via a stabilization and activation of p53 (Wang *et al.*, *Oncogene* 19:1843-48 (2000)); and, Shou *et al.*, *Oncogene* 21:878-89 (2002)). Additionally, Fedi *et al.* (1999) purportedly showed that Dkk-1 can block the Wnt2-induced oncogenic transformation of NIH-3T3 cells.

Furthermore, it has been suggested that Dkk expression can be modulated by BMP signaling in the developing skeleton (Mukhopadhyay *et al.*, *Dev. Cell.* 1:423-34 (2001); and Grotewold *et al.*, *EMBO J.* 21:966-75 (2002)). Grotewold et al. additionally describe altered Dkk expression levels in response to stress signals including UV irradiation and other genotoxic stimuli. They propose that Dkk expression is pro-apoptotic. In animals expressing HBM constructs conferring high bone mass, a reduced osteoblast apoptosis effect was observed. Thus, HBM and HBM-like variants may control/alter Dkk's role in programmed cell death. Other agents which potentially modulate Dkk activity include the Dkk interacting proteins identified in Figure 5.

In one embodiment, the relative amounts of Dkk or a Dkk interacting protein of a cell population that has been exposed to the agent to be tested is compared to an un-exposed control cell population. Antibodies can be used to monitor the differential expression of the protein in the different cell populations. Cell lines or populations are exposed to the agent to be tested under appropriate conditions and time. Cellular lysates may be prepared from the exposed cell line or population and a control, unexposed cell line or population. The cellular lysates are then analyzed with the probe, as would be known in the art. *See, e.g.,* Ed Harlow and David Lane, Antibodies: A Laboratory Manual (Cold Spring Harbor, NY, 1988) and Ed Harlow and David Lane, Using Antibodies: A Laboratory Manual (Cold Spring Harbor, NY 1998).

For example, N- and C- terminal fragments of Dkk can be expressed in bacteria and used to search for proteins which bind to these fragments. Fusion proteins, such as His-tag or GST fusion to the N- or C-terminal regions of Dkk (or to biologically active domains of Dkk-1) or a whole Dkk protein can be prepared. These fusion proteins can be coupled to, for example, Talon or Glutathione-Sepharose beads and then probed with cell lysates to identify molecules which bind to Dkk. Prior to lysis, the cells may be treated with purified Wnt proteins, RNA, or drugs which may modulate Wnt signaling or proteins that interact with downstream elements of the Wnt pathway. Lysate proteins binding to the fusion proteins can be resolved by SDS-PAGE, isolated and identified by, for example protein sequencing or mass spectroscopy, as is known in the art. *See, e.g.*, Protein Purification Applications: A Practical Approach (Simon Roe, ed., 2nd ed.

Oxford Univ. Press, 2001) and "Guide to Protein Purification" in *Meth. Enzymology* vol. 182 (Academic Press, 1997).

The activity of Dkk, a Dkk interacting protein, or a complex of Dkk with LRP5/LRP6/HBM may be affected by compounds which modulate the interaction between Dkk and a Dkk interacting protein (such as those shown in Figure 5) and/or Dkk and LRP5/LRP6/HBM. The present invention provides methods and research tools for the discovery and characterization of these compounds. The interaction between Dkk and a Dkk interacting protein and/or Dkk and LRP5/6/HBM may be monitored *in vivo* and *in vitro*. Compounds which modulate the stability of a Dkk - LRP5/LRP6/HBM complex are potential therapeutic compounds. Example *in vitro* methods include: Binding LRP5/6/HBM, Dkk, or a Dkk interacting protein to a sensor chip designed for an instrument such are made by Biacore (Uppsala, Sweden) for the performance of an plasmon resonance spectroscopy observation. In this method, the chip with one of Dkk, a Dkk interacting protein, or LRP5/6 is first exposed to the other under conditions which permit them to form the complex. A test compound is then introduced and the output signal of the instrument provides an indication of any effect exerted by the test compound. By this method, compounds may be rapidly screened. Another, in vitro, method is exemplified by the SAR-by-NMR methods (Shuker *et al.*, *Science*. 274:1531-4 (1996)). Briefly, a Dkk-1 binding domain and/or LRP 5 or 6 LBD are expressed and purified as ^{15}N labeled protein by expression in labeled media. The labeled protein(s) are allowed to form the complex in solution in an NMR sample tube. The heteronuclear correlation spectrum in the presence and absence of a test compound provides data at the level of individual residues with regard to interactions with the test compound and changes at the protein-protein interface of the complex. One of skill in the art knows of many other protocols, e.g. affinity capillary electrophoresis (Okun *et al. J Biol Chem* 276:1057-62 (2001); Vergun and Chu, *Methods,* 19:270-7 (1999)), fluorescence spectroscopy, electron paramagnetic resonance, etc. which can monitor the modulation of a complex and/or measure binding affinities for complex formation.

In vitro protocols for monitoring the modulation of a Dkk/LRP5/LRP6/HBM complex include the yeast two hybrid protocol. The yeast two hybrid method may be used to monitor the modulation of a complex in vivo by monitoring the expression of genes activated by the formation of a complex of fusion proteins of Dkk and LRP ligand binding domains. Nucleic acids according to the invention which encode the interacting Dkk and LRP LBD domains are incorporated into bait and prey plasmids. The Y2H protocol is performed in the presence of one or more test compounds. The modulation of the complex is observed by a change in expression of the complex activated gene. It will be appreciated by one skilled in the art that test compounds can be added to the assay directly or, in the case of proteins, can be coexpressed in the yeast with the bait and prey compounds. Similarly, fusion proteins of Dkk and Dkk interacting proteins can also be used in a Y2H screen to identify other proteins which modulate the Dkk/Dkk interacting protein complex.

Assay protocols such as these may be used in methods to screen for compounds, drugs, treatments which modulate the Dkk/Dkk interacting protein and/or Dkk/LRP5/6 complex, whether such modulation occurs by competitive binding, or by altering the structure of either LRP 5/6 or Dkk at the binding site, or by stabilizing or destablizing the protein-protein interface. It may be anticipated that peptide aptamers may competitively bind, although induction of an altered binding site structure by steric effects is also possible.

12.1 Antibodies and Antibody Fragments

Polyclonal and monoclonal antibodies and fragments of these antibodies which bind to Dkk or LRP5/LRP6/HBM can be prepared as would be known in the art. For example, suitable host animals can be immunized using appropriate immunization protocols and the peptides, polypeptides or proteins of the invention. Peptides for use in immunization are typically about 8-40 residues long. If necessary or desired, the polypeptide immunogens can be conjugated to suitable carriers. Methods for preparing immunogenic conjugates with carriers such as bovine serum albumin (BSA), keyhole limpet hemocyanin (KLH), or other carrier proteins are well known in the art (*See*,

Harlow *et al.*, 1988). In some circumstances, direct conjugation using, for example, carbodiimide reagents, may be effective; in other instances linking reagents such as those supplied by Pierce Chemical Co., Rockford, IL, may be desirable to provide accessibility to the polypeptide or hapten. The hapten peptides can be extended at either the amino or carboxy terminus with a cysteine residue or interspersed with cysteine residues, for example, to facilitate linking to a carrier. Administration of the immunogens is conducted generally by injection over a suitable time period and with use of suitable adjuvants, as is generally understood in the art. During the immunization schedule, titers of antibodies are taken to determine adequacy of antibody formation.

Anti-peptide antibodies can be generated using synthetic peptides, for example, the peptides derived from the sequence of any Dkk, including Dkk-1, or LRP5/LRP6/HBM. Synthetic peptides can be as small as 2-3 amino acids in length, but are preferably at least 3, 5, 10, or 15 or more amino acid residues long. Such peptides can be determined using programs such as DNAStar. The peptides are coupled to KLH using standard methods and can be immunized into animals such as rabbits. Polyclonal anti-Dkk or anti-LRP5/LRP6/HBM peptide antibodies can then be purified, for example using Actigel beads containing the covalently bound peptide.

While the polyclonal antisera produced in this way may be satisfactory for some applications, for pharmaceutical compositions, use of monoclonal preparations is preferred. Immortalized cell lines which secrete the desired monoclonal antibodies may be prepared using the standard method of Kohler and Milstein or modifications which effect immortalization of lymphocytes or spleen cells, as is generally known (See, e.g., Harlow *et al.*, 1988 and 1998). The immortalized cell lines secreting the desired antibodies can be screened by immunoassay in which the antigen is the peptide hapten, polypeptide or protein. When the appropriate immortalized cell culture secreting the desired antibody is identified, the cells can be cultured either *in vitro* or by production in ascites fluid.

The desired monoclonal antibodies are then recovered from the culture supernatant or from the ascites supernatant. Fragments of the monoclonal antibodies

which contain the immunologically significant portion can be used as agonists or antagonists of Dkk activity. Use of immunologically reactive fragments, such as the Fab, scFV, Fab', of $F(ab')_2$ fragments are often preferable, especially in a therapeutic context, as these fragments are generally less immunogenic than the whole immunoglobulin.

The antibodies or fragments may also be produced, using current technology, by recombinant means. Regions that bind specifically to the desired regions of Dkk or LRP5/LRP6/HBM can also be produced in the context of chimeras with multiple species origin. Antibody reagents so created are contemplated for use diagnostically or as stimulants or inhibitors of Dkk activity.

In one embodiment, antibodies against Dkk, bind Dkk with high affinity, i.e., ranging from 10^{-5} to 10^{-9} M. Preferably, the anti-Dkk antibody will comprise a chimeric, primate, Primatized®, human or humanized antibody. Also, the invention embraces the use of antibody fragments, *e.g.*, Fab's, Fv's, Fab's, $F(ab)_2$, and aggregates thereof.

Another embodiment contemplates chimeric antibodies which recognize Dkk or LRP5/LRP6/HBM. A chimeric antibody is intended to refer to an antibody with non-human variable regions and human constant regions, most typically rodent variable regions and human constant regions.

A "primatized® antibody" refers to an antibody with primate variable regions, *e.g.*, CDR's, and human constant regions. Preferably, such primate variable regions are derived from an Old World monkey.

A "humanized antibody" refers to an antibody with substantially human framework and constant regions, and non-human complementarity-determining regions (CDRs). "Substantially" refers to the fact that humanized antibodies typically retain at least several donor framework residues (*i.e.*, of non-human parent antibody from which CDRs are derived).

Methods for producing chimeric, primate, primatized®, humanized and human antibodies are well known in the art. See, *e.g.*, U.S. Patent 5,530,101, issued to Queen *et al.*; U.S. Patent 5,225,539, issued to Winter *et al.*; U.S. Patents 4,816,397 and

4,816,567, issued to Boss *et al.* and Cabilly *et al.* respectively, all of which are incorporated by reference in their entirety.

The selection of human constant regions may be significant to the therapeutic efficacy of the subject anti-Dkk or LRP5/LRP6/HBM antibody. In a preferred embodiment, the subject anti-Dkk or LRP5/LRP6/HBM antibody will comprise human, gamma 1, or gamma 3 constant regions and, more preferably, human gamma 1 constant regions.

Methods for making human antibodies are also known and include, by way of example, production in SCID mice, and *in vitro* immunization.

The subject anti-Dkk or LRP5/LRP6/HBM antibodies can be administered by various routes of administration, typically parenteral. This is intended to include intravenous, intramuscular, subcutaneous, rectal, vaginal, and administration with intravenous infusion being preferred.

The anti-Dkk or LRP5/LRP6/HBM antibody will be formulated for therapeutic usage by standard methods, *e.g.*, by addition of pharmaceutically acceptable buffers, *e.g.*, sterile saline, sterile buffered water, propylene glycol, and combinations thereof.

Effective dosages will depend on the specific antibody, condition of the patient, age, weight, or any other treatments, among other factors. Typically effective dosages will range from about 0.001 to about 30 mg/kg body weight, more preferably from about 0.01 to 25 mg/kg body weight, and most preferably from about 0.1 to about 20 mg/kg body weight.

Such administration may be effected by various protocols, *e.g.*, weekly, bi-weekly, or monthly, depending on the dosage administered and patient response. Also, it may be desirable to combine such administration with other treatments.

Antibodies to Dkk-1 interacting proteins, such as those identified in Figure 5, are also contemplated according to the present invention, and can be used similarly to the Dkk-1 antibodies mentioned in the above methodology.

The antibodies of the present invention can be utilized in experimental screening, as diagnostic reagents, and in therapeutic compositions.

12.2 Chemical Libraries

Agents that are assayed by these methods can be randomly selected or rationally selected or designed. As used herein, an agent is said to be randomly selected when the agent is chosen randomly without considering the specific sequences involved in the association of Dkk-1 alone, Dkk-1 interacting proteins alone, or with their associated substrates, binding partners, etc. An example of randomly selected agents is the use of a chemical library or a peptide combinatorial library, or a growth broth of an organism.

The agents of the present invention can be, as examples, peptides, small molecules, vitamin derivatives, as well as carbohydrates. A skilled artisan can readily recognize that there is no limit as to the structural nature of the agents of the present invention.

12.3 Peptide Synthesis

The peptide agents of the invention can be prepared using standard solid phase (or solution phase) peptide synthesis methods, as is known in the art. In addition, the DNA encoding these peptides may be synthesized using commercially available oligonucleotide synthesis instrumentation and produced recombinantly using standard recombinant production systems. The production of polypeptides using solid phase peptide synthesis is necessitated if non-nucleic acid-encoded amino acids are to be included.

13. Uses for Agents that Modulate at Least One Activity of Dkk, a Dkk Interacting Protein, a Dkk/Dkk Interacting Protein Complex, or a Dkk/LRP5 or Dkk/LRP6 Complex

The proteins and nucleic acids of the invention, such as the proteins or polypeptides containing an amino acid sequence of LRP5, Dkk, and Dkk interacting proteins are involved in bone mass modulation and lipid modulation of other Wnt pathway mediated activity. Agents that modulate (*i.e.*, up and down-regulate) the expression of Dkk or Dkk interacting proteins, or agents, such as agonists and

antagonists respectively, of at least one activity of Dkk or a Dkk interacting protein may be used to modulate biological and pathologic processes associated with the function and activity of Dkk or a Dkk interacting protein.

As used herein, a subject can be preferably any mammal, so long as the mammal is in need of modulation of a pathological or biological process modulated by a protein of the invention. The term "mammal" means an individual belonging to the class *Mammalia*. The invention is particularly useful in the treatment of human subjects.

As used herein, a biological or pathological process modulated by Dkk or a Dkk interacting protein may include binding of Dkk to a Dkk interacting protein, Dkk to LRP5 or LRP6 or release therefrom, inhibiting or activating Dkk or a Dkk interacting protein mRNA synthesis or inhibiting Dkk or Dkk interacting protein modulated inhibition of LRP5 or LRP6 mediated Wnt signaling. Further bone-related markers may be observed such as alkaline phosphatase activity, osteocalcin production, or mineralization.

Pathological processes refer to a category of biological processes which produce a deleterious effect. For example, expression or up-regulation of expression of LRP5 or LRP6 and/or Dkk and/or a Dkk interacting protein may be associated with certain diseases or pathological conditions. As used herein, an agent is said to modulate a pathological process when the agent statistically significantly ($p < 0.05$) alters the process from its base level in the subject. For example, the agent may reduce the degree or severity of the process mediated by that protein in the subject to which the agent was administered. For instance, a disease or pathological condition may be prevented, or disease progression modulated by the administration of agents which reduce or modulate in some way the expression or at least one activity of a protein of the invention.

As LRP5/6 and Dkk are involved both directly and indirectly in bone mass modulation, one embodiment of this invention is to use Dkk or Dkk interacting protein expression as a method of diagnosing a bone condition or disease. Certain markers are associated with specific Wnt signaling conditions (*e.g., TCF/LEF* activation). Diagnostic tests for bone conditions may include the steps of testing a sample or an

extract thereof for the presence of Dkk or Dkk interacting protein nucleic acids (*i.e.*, DNA or RNA), oligomers or fragments thereof or protein products of TCF/LEF regulated expression. For example, standard *in situ* hybridization or other imaging techniques can be utilized to observe products of Wnt signaling.

This invention also relates to methods of modulating bone development or bone loss conditions. Inhibition of bone loss may be achieved by inhibiting or modulating changes in the LRP5/6 mediated Wnt signaling pathway. For example, absence of LRP5 activity may be associated with low bone mass. Increased activity LRP5 may be associated with high bone mass. Therefore, modulation of LRP5 activity will in turn modulate bone development. Modulation of the Dkk/LRP5/6 or Dkk/Dkk interacting protein complex via agonists and antagonists is one embodiment of a method to regulate bone development. Such modulation of bone development can result from inhibition of the activity of, for example, a Dkk/LRP(5/6) protein complex, a Dkk/Dkk interacting protein complex, upregulated transcription of the *LRP5* gene or inhibited translation of Dkk or Dkk interacting protein mRNA.

The agents of the present invention can be provided alone, or in combination with other agents that modulate a particular pathological process. As used herein, two agents are said to be administered in combination when the two agents are administered simultaneously or are administered independently in a fashion such that the agents will act at the same time.

The agents of the present invention can be administered via parenteral, subcutaneous (sc), intravenous (iv), intramuscular (im), intraperitoneal (ip), transdermal or buccal routes. Alternatively, or concurrently, administration may be by the oral route. The dosage administered will be dependent upon the age, health, and weight of the recipient, kind of concurrent treatment, if any, frequency of treatment, and the nature of the effect desired.

The present invention further provides compositions containing one or more agents which modulate expression or at least one activity of a protein of the invention. While individual needs vary, determination of optimal ranges of effective amounts of each component is within the skill of the art. Typical dosages of the active agent which

mediate Dkk or Dkk interacting protein activity comprise from about 0.0001 to about 50 mg/kg body weight. The preferred dosages comprise from about 0.001 to about 50 mg/kg body weight. The most preferred dosages comprise from about 0.1 to about 1 mg/kg body weight. In an average human of 70 kg, the range would be from about 7 μg to about 3.5 g, with a preferred range of about 0.5 mg to about 5 mg.

In addition to the pharmacologically active agent, the compositions of the present invention may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically for delivery to the site of action. Suitable formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form, for example, water-soluble salts. In addition, suspensions of the active compounds as appropriate oily injection suspensions may be administered. Suitable lipophilic solvents or vehicles include fatty oils, for example, sesame oil, or synthetic fatty acid esters, (*e.g.*, ethyl oleate or triglycerides). Aqueous injection suspensions may contain substances which increase the viscosity of the suspension include, for example, sodium carboxymethyl cellulose, sorbitol and/or dextran. Optionally, the suspension may also contain stabilizers. Liposomes and other non-viral vectors can also be used to encapsulate the agent for delivery into the cell.

The pharmaceutical formulation for systemic administration according to the invention may be formulated for enteral, parenteral, or topical (top) administration. Indeed, all three types of formulations may be used simultaneously to achieve systemic administration of the active ingredient.

Suitable formulations for oral administration include hard or soft gelatin capsules, pills, tablets, including coated tablets, elixirs, suspensions, syrups or inhalations and controlled release forms thereof.

Potentially, any compound which binds Dkk or a Dkk interacting protein or modulates the Dkk/LRP5 or Dkk/LRP6 or Dkk/Dkk interacting protein complex may be a therapeutic compound. In one embodiment of the invention, a peptide or nucleic acid aptamer according to the invention is used in a therapeutic composition. Such compositions may comprise an aptamer, or a LRP5 or LRP6 fragment unmodified or

modified. In another embodiment, the therapeutic compound comprises a Dkk-1 interacting protein, or biologically active fragment thereof.

Nucleic acid aptamers have been used in compositions for example by chemical bonding to a carrier molecule such as polyethylene glycol (PEG) which may facilitate uptake or stabilize the aptamer. A di-alkylgylcerol moiety attached to an RNA will embed the aptamer in liposomes, thus stabilizing the compound. Incorporating chemical substitutions (i.e. changing the 2'OH group of ribose to a 2'NH in RNA confers ribonuclease resistance) and capping, etc. can prevent breakdown. Several such techniques are discussed for RNA aptamers in Brody and Gold *(Rev. Mol. Biol.* 74:3-13 (2000)).

Peptide aptamers may by used in therapeutic applications by the introduction of an expression vector directing aptamer expression into the affected tissue such as for example by retroviral delivery, by encapsulating the DNA in a delivery complex or simple by naked DNA injection. Or, the aptamer itself or a synthetic analog may be used directly as a drug. Encapsulation in polymers and lipids may assist in delivery. The use of peptide aptamers as therapeutic and diagnostic agents is reviewed by Hoppe-Syler and Butz *(J. Mol. Med.* 78:426-430 (2000)).

In another aspect of the invention. The structure of a constrained peptide aptamer of the invention may be determined such as by NMR or X-ray crystallography. (Cavanagh et al., Protein NMR Spectroscopy: Principles and Practice, Academic Press, 1996; Drenth, Principles of Protein X-Ray Crystallography, Springer Verlag, 1999) Preferably the structure is determined in complex with the target protein. A small molecule analog is then designed according to the positions of functional elements of the 3D structure of the aptamer. (Guidebook on Molecular Modeling in Drug Design, Cohen, Ed., Academic Press, 1996; Molecular Modeling and Drug Design (Topics in Molecular and Structural Biology), Vinter and Gardner Eds., CRC Press, 1994) Thus the present invention provides a method for the design of effective and specific drugs which modulate the activity of Dkk, Dkk interacting proteins, Dkk/Dkk interacting protein complex and the Dkk/LRP complex. Small molecule mimetics of the peptide aptamers of the present invention are encompassed within the scope of the invention.

In practicing the methods of this invention, the compounds of this invention may be used alone or in combination, or in combination with other therapeutic or diagnostic agents. In certain preferred embodiments, the compounds of this invention may be co-administered along with other compounds typically prescribed for these conditions according to generally accepted medical practice. For example, the compounds of this invention can be administered in combination with other therapeutic agents for the treatment of bone loss. Bone loss mediating agents include bone resorption inhibitors such as bisphosphonates (*e.g.*, alendronic acid, clodronic acid, etidronic acid, pamidronic acid, risedronic acid and tiludronic acid), vitamin D and vitamin D analogs, cathepsin K inhibitors, hormonal agents (*e.g.*, calcitonin and estrogen), and selective estrogen receptor modulators or SERMs (e.g., raloxifene). And bone forming agents such as parathyroid hormone (PTH) and bone morphogenetic proteins (BMP).

Additionally contemplated are combinations of agents which regulate Dkk-1 and agents which regulate lipid levels such as HMG-CoA reductase inhibitors (*i.e.*, statins such as Mevacor®, Lipitor® and other inhibitors such as Baycol®, Lescol®, Pravachol® and Zocor®), bile acid sequestrants (e.g., Colestid® and Welchol®), fibric acid derivatives (Atromid-S®, Lopid®, Tricor®), and nicotinic acid.

[0001] The compounds of this invention can be utilized *in vivo*, ordinarily in vertebrates and preferably in mammals, such as humans, sheep, horses, cattle, pigs, dogs, cats, rats and mice, or *in vitro*.

14. <u>Transgenic Animals</u>

Transgenic animal models can be created which conditionally express Dkk and/or LRP5 or LRP6 and/or Dkk interacting proteins, such as those shown in Figure 5. These animals can be used as research tools for the study of the physiological effects of the Dkk-1/Dkk-1 interacting protein interaction and/or the LRP5 / Dkk interaction. Alternatively, transgenic animals can be created which express a transgenic form of Dkk alone or in addition to a transgenic form of HBM or express Dkk interacting proteins alone or in addition to a transgenic form of Dkk. Transgenic animals expressing HBM or LRP5 can be crossed with transgenic animals expressing Dkk or

Dkk interacting proteins to obtain heterozygote as well as homozygote animals which express both desired genes.

Animal models may be created to directly modulate the Dkk/Dkk interacting protein or Dkk/ LRP5 interaction activity *in vivo* to serve as a research tool for determining the efficacy of candidate compounds which modulate the Dkk/Dkk interacting protein or LRP5 / Dkk interaction activity *in vitro*. Animals, such as transgenic mice, can be created using the techniques employed to make transgenic mice that express for example, human Dkk or a Dkk interacting protein, or knockouts (KO), which may be conditional, of the gene encoding mouse Dkk or Dkk interacting protein. Knock-in animals include animals wherein genes have been introduced and animals wherein a gene that was previously knocked-out is reintroduced into the animal. Other transgenic animals can be created with inducible forms of Dkk or a Dkk interacting protein to study the effects of the gene on bone mass development and loss as well as lipid level regulation. These animals can also be used to study long term effects of Dkk or Dkk interacting protein modulation. Transgenic animals may be created to express peptide aptamers, or produce RNA aptamers. The transgenic vectors may direct expression in a tissue specific manner by the use of tissue specific promoters. In a preferred embodiment, a peptide aptamer fusion protein is expressed using a bone specific promoter. Such systems can provide a tissue specific knock-out of Dkk or Dkk interacting protein activity.

General methods for creating transgenic animals are known in the art, and are described in, for example, Strategies in Transgenic Animal Science (Glenn M. Monastersky and James M. Robl eds., ASM Press; Washington, DC, 1995); Transgenic Animal Technology: A Laboratory Handbook (Carl A. Pinkert ed., Academic Press 1994); Transgenic Animals (Louis Marie Houdebine, ed., Harwood Academic Press, 1997); Overexpression and Knockout of Cytokines in Transgenic Mice (Chaim O. Jacob, ed., Academic Press 1994); Microinjection and Transgenesis: Strategies and Protocols (Springer Lab Manual) (Angel Cid-Arregui and Alejandro Garcia-Carranca, eds., Springer Verlag 1998); and Manipulating the Mouse Embryo: A Laboratory Manual (Brigid Hogan *et al.*, eds., Cold Spring Harbor Laboratory Press 1994).

15. Peptide and Nucleotide Aptamers and Peptide Aptamer Mimetics

Another embodiment contemplates the use of peptide and nucleotide aptamer technology to screen for agents which interact with Dkk, which block Dkk from interacting with LRP5 or LRP6, or which block any other Dkk ligand interaction, or which interact with Dkk interacting proteins, such as those shown in Figure 5. Peptide aptamers are molecules in which a variable peptide domain is displayed from a scaffold protein. Thioredoxin A (trxA) is commonly used for a scaffold. The peptide insert destroys the catalytic site of trxA. It is recognized that numerous proteins may also be used as scaffolding proteins to constrain and/or present a peptide aptamer. Other scaffold proteins that could display a constrained peptide aptamer could include staphylococcal nuclease, the protease inhibitor eglin C, the *Streptomyces tendea* alpha-amylase inhibitor Tendamistat, Sp1, and green fluorescent protein (GFP) (reviewed in Hoppe-Seyler *et al.*, *J. Steroid Biochem Mol. Biol.* 78:105-11 (2001)), and the S1 nuclease from *Staphylococcus* or M13 for phage display. Any molecule to which the aptamer could be anchored and presented in its bioactive conformation would be suitable.

Aptamers can then specifically bind to a given target protein *in vitro* and *in vivo* and have the potential to selectively block the function of their target protein. Peptide aptamers are selected from randomized expression libraries on the basis of their *in vivo* binding capacity to the desired target protein. Briefly, a target protein (e.g., Dkk, a Dkk interacting protein, or LRP5/6) is linked to a heterologous DNA binding domain (BD) and expressed as bait in a yeast test strain. Concomitantly, a library coding for different peptides (e.g., 16-mers) of randomized sequence inserted in a scaffold protein sequence, which are linked to a heterologous transcriptional activation domain (AD) is expressed as prey. If a peptide binds to a target protein, a functional transcription factor is reconstituted, in which the BD and AD are bridged together by interacting proteins. This transcription factor is then able to activate the promoter of a marker gene which can be monitored by colorimetric enzymatic assays or by growth selection. Additional variation, methods of preparing and screening methodologies are described in, for example, Hoppe-Seyler *et al.*, *J. Mol. Med.* 78: 426-430 (2000).

Nucleotide aptamers are described for example in Brody *et al., Trends Mol. Biotechnol.* 74: 5-13 (2000). Additional methods of making and using nucleotide aptamers include SELEX, *i.e.*, Systematic Evolution of Ligands by Exponential Enrichment. SELEX is a process of isolating oligonucleotide ligands of a chosen target molecule (see Tuerk and Gold, *Science* 249:505-510 (1990); U.S. Pat. Nos. 5,475,096, 5,595,877, and 5,660,985). SELEX, as described in Tuerk and Gold, involves admixing the target molecule with a pool of oligonucleotides (*e.g.*, RNA) of diverse sequences; retaining complexes formed between the target and oligonucleotides; recovering the oligonucleotides bound to the target; reverse-transcribing the RNA into DNA; amplifying the DNA with polymerase chain reactions (PCR); transcribing the amplified DNA into RNA; and repeating the cycle with ever increasing binding stringency. Three enzymatic reactions are required for each cycle. It usually takes 12-15 cycles to isolate aptamers of high affinity and specificity to the target. An aptamer is an oligonucleotide that is capable of binding to an intended target substance but not other molecules under the same conditions.

In another reference, Bock *et al.*, *Nature* 355:564-566 (1990), describe a different process from the SELEX method of Tuerk and Gold in that only one enzymatic reaction is required for each cycle (*i.e.*, PCR) because the nucleic acid library in Bock's method is comprised of DNA instead of RNA. The identification and isolation of aptamers of high specificity and affinity with the method of Bock *et al.* still requires repeated cycles in a chromatographic column.

Other nucleotide aptamer methods include those described by Conrad *et al.*, *Meth. Enzymol.* 267:336-367 (1996). Conrad *et al.* describe a variety of methods for isolating aptamers, all of which employ repeated cycles to enrich target-bound ligands and require a large amount of purified target molecules. More recently described methods of making and using nucleotide aptamers include, but are not limited to those described in U.S. Patent Nos. 6,180,348; 6,051,388; 5,840,867; 5,780,610, 5,756,291 and 5,582,981.

Potentially, any compound which binds Dkk or a Dkk interacting protein or modulates the Dkk/Dkk interacting protein or Dkk/LRP5 or Dkk/LRP6 complex may be

a therapeutic compound. In one embodiment of the invention, a peptide or nucleic acid aptamer according to the invention is used in a therapeutic composition. Such compositions may comprise an aptamer, or a LRP5 or LRP6 fragment unmodified or modified.

Nucleic acid aptamers have been used in compositions for example by chemical bonding to a carrier molecule such as polyethylene glycol (PEG) which may facilitate uptake or stabilize the aptamer. A di-alkylglycerol moiety attached to an RNA will embed the aptamer in liposomes, thus stabilizing the compound. Incorporating chemical substitutions (*i.e.*, changing the 2'-OH group of ribose to a 2'-NH in RNA confers ribonuclease resistance) and capping, etc. can prevent breakdown. Several such techniques are discussed for RNA aptamers in Brody and Gold *Rev. Mol. Biol.* 74:3-13 (2000).

Peptide aptamers may by used in therapeutic applications by the introduction of an expression vector directing aptamer expression into the affected tissue such as for example by retroviral delivery, by encapsulating the DNA in a delivery complex or simple by naked DNA injection. Or, the aptamer itself or a synthetic analog may be used directly as a drug. Encapsulation in polymers and lipids may assist in delivery. The use of peptide aptamers as therapeutic and diagnostic agents is reviewed by Hoppe-Syler and Butz *J. Mol. Med.* 78:426-430 (2000).

In another aspect of the invention, the structure of a constrained peptide aptamer of the invention may be determined such as by NMR or X-ray crystallography. (Cavanagh et al., Protein NMR Spectroscopy : Principles and Practice, Academic Press, 1996; Drenth, Principles of Protein X-Ray Crystallography, Springer Verlag, 1999) Preferably the structure is determined in complex with the target protein. A small molecule analog is then designed according to the positions of functional elements of the 3D structure of the aptamer. (Guidebook on Molecular Modeling in Drug Design, Cohen, Ed., Academic Press, 1996; Molecular Modeling and Drug Design (Topics in Molecular and Structural Biology), Vinter and Gardner Eds., CRC Press, 1994) Thus, a method is provided for the design of effective and specific drugs which modulate the activity of Dkk, Dkk interacting proteins, Dkk/Dkk interacting protein

complex, and the Dkk/LRP complex. Small molecule mimics of the peptide aptamers of the present invention are also encompassed within the scope of the invention.

16. Alternative Variants of LRP5/LRP6 Having HBM Activity

A structural model of the LRP5/Zmax1 first beta-propeller module was generated based on a model prediction in Springer et al., (1998) *J. Molecular Biology,* 283:837-862. Based on the model, certain amino acid residues were identified as important variants of LRP5/HBM/Zmax1. The following three categories provide examples of such variants:

The shape of the beta-propeller resembles a disk with inward-sloping sides and a hole down the middle. Residue 171 is in a loop on the outer or top surface of the domain in blade 4 of propeller module 1. Thus, variants comprising changed residues in structurally equivalent positions in other blades; as well as residues that are slightly more interior to the binding pocket, but still accessible to the surface, are important embodiments of the present invention for the study of bone mass modulation by LRP5/HBM, for the development of pharmaceuticals and treatments of bone mass disorders, and for other objectives of the present invention. The following are examples of such variants:

A214V (a position equivalent to 171 in blade 5; alanine is not conserved in other propellers),

E128V (a position equivalent to 171 in blade 3; glutamate is not conserved in other propellers),

A65V (a position equivalent to 171 in blade 2; alanine is conserved in propellers 1-3 but not 4),

G199V (an accessible interior position in blade 5; glycine is conserved in propellers 1-3 but not 4), and

M282V (accessible interior position in blade 1; methionine is conserved in propellers 1-3 but not 4).

LRP5/Zmax1 has four beta-propeller structures; the first three beta-propeller modules conserve a glycine in the position corresponding to residue 171 in human

LRP5/Zmax1. Therefore, variants bearing a valine in the equivalent positions in the other propellers are important embodiments of the present invention. The following variants are useful for the study of bone mass modulation by LRP5/HBM, for the development of pharmaceuticals and treatments of bone mass disorders, and for other objectives of the present invention: G479V, G781V, and Q1087V.

The G171V HBM polymorphism results in "occupied space" of the beta-propeller 1, with the side-chain from the valine residue sticking out into an open binding pocket and potentially altering a ligand/protein interaction. The glycine residue is conserved in LRP5/Zmax1 propellers 1, 2 and 3 but is a glutamine in propeller 4. Therefore, the following variants of LRP5/HBM are important embodiments of the present invention for the study of bone mass modulation by LRP5/HBM, for the development of pharmaceuticals and treatments of bone mass disorders, and for other objectives of the present invention:

G171K (which introduces a charged side-chain),
G171F (which introduces a ringed side-chain),
G171I (which introduces a branched side-chain), and
G171Q (which introduces the propeller 4 residue).

Furthermore, LRP6 is the closest homolog of LRP5/Zmax1. LRP6 has a beta-propeller structure predicted to be similar, if not identical to Zmax1. The position corresponding to glycine 171 in human LRP5/Zmax1 is glycine 158 of human LRP6. Thus, corresponding variants of LRP6 are an important embodiment of the present invention for the study of the specificity of LRP5/Zmax1 versus its related family member, for the development of pharmaceuticals and treatments of bone mass disorders, and for other objectives of the present invention. Specifically, for example, a glycine to valine substitution at the structurally equivalent position, residue 158, of human LRP6 and similar variants of other species' LRP6 homologs represent important research tools.

Site-directed mutants of LRP5 were generated in the full-length human LRP5 cDNA using the QuikChange XL-Site-Directed Mutagenesis Kit (catalog #200516,

Stratagene, La Jolla, CA) following the manufacturer's protocol. The mutant sequences were introduced using complementary synthetic oligonucleotides:

A65V: TGGTCAGCGGCCTGGAGGATGTGGCCGCAGTGGACTTCC (SEQ ID NO:129) and GGAAGTCCACTGCGGCCACATCCTCCAGGCCGCTGACCA (SEQ ID NO:130)

E128V: AAGCTGTACTGGACGGACTCAGTGACCAACCGCATCGAGG (SEQ ID NO:131) and CCTCGATGCGGTTGGTCACTGAGTCCGTCCAGTACAGCTT (SEQ ID NO:132)

G171K: ATGTACTGGACAGACTGGAAGGAGACGCCCCGGATTGAGCG (SEQ ID NO: 133) and CGCTCAATCCGGGGCGTCTCCTTCCAGTCTGTCCAGTACAT (SEQ ID NO:134)

G171F: ATGTACTGGACAGACTGGTTTGAGACGCCCCGGATTGAGCG (SEQ ID NO:135) and CGCTCAATCCGGGGCGTCTCAAACCAGTCTGTCCAGTACAT (SEQ ID NO:136)

G171I: ATGTACTGGACAGACTGGATTGAGACGCCCCGGATTGAGCG (SEQ ID NO:137) and CGCTCAATCCGGGGCGTCTCAATCCAGTCTGTCCAGTACAT (SEQ ID NO:138)

G171Q: ATGTACTGGACAGACTGGCAGGAGACGCCCCGGATTGAGCG (SEQ ID NO:139) and CGCTCAATCCGGGGCGTCTCCTGCCAGTCTGTCCAGTACAT (SEQ ID NO:140)

G199V: CGGACATTTACTGGCCCAATGTACTGACCATCGACCTGGAGG (SEQ ID NO:141) and CCTCCAGGTCGATGGTCAGTACATTGGGCCAGTAAATGTCCG (SEQ ID NO:142)

A214V: AGCTCTACTGGGCTGACGTCAAGCTCAGCTTCATCCACCG (SEQ ID NO: 143) and CGGTGGATGAAGCTGAGCTTGACGTCAGCCCAGTAGAGCT (SEQ ID NO:144)

M282V: GAGTGCCCTCTACTCACCCGTGGACATCCAGGTGCTGAGCC (SEQ ID NO:145) and GGCTCAGCACCTGGATGTCCACGGGTGAGTAGAGGGCACTC (SEQ ID NO:146)

G479V: CATGTACTGGACAGACTGGGTAGAGAACCCTAAAATCGAGTGTGC (SEQ ID NO:147) and GCACACTCGATTTTAGGGTTCTCTACCCAGTCTGTCCAGTACATG (SEQ ID NO:148)

G781V: CATCTACTGGACCGAGTGGGTCGGCAAGCCGAGGATCGTGCG (SEQ ID NO:149) and CGCACGATCCTCGGCTTGCCGACCCACTCGGTCCAGTAGATG (SEQ ID NO:150)

Q1087V: GTACTTCACCAACATGGTGGACCGGGCAGCCAAGATCGAACG (SEQ ID NO:151) and CGTTCGATCTTGGCTGCCCGGTCCACCATGTTGGTGAAGTAC (SEQ ID NO:152)

LRP6 G158V: GTACTGGACAGACTGGGTAGAAGTGCCAAAGATAGAACGTGC (SEQ ID NO:153) and GCACGTTCTATCTTTGGCACTTCTACCCAGTCTGTCCAGTAC (SEQ ID NO:154).

All constructs were sequence verified to ensure that only the engineered modification was present in the gene. Once verified, each variant was functionally evaluated in the TCF-luciferase assay in U2OS cells (essentially as described in Example 7. Other functional evaluations could also be performed, such as the Xenopus embryo assay (essentially as described in Example 6), or other assays to evaluate Wnt

signaling, Dkk modulation, or anabolic bone effect. Binding of these mutants to Dkk, LRP-interacting proteins, Dkk-interacting proteins, or peptide aptamers to any of the preceding could also be investigated in a variety of ways such as in a two-hybrid system (such as in yeast as described in this application), or other methods.

Figure 24 shows the effects of the G171F mutation in propeller 1 of LRP5. This mutation is at the same position as HBM's G171V substitution. Expression of G171F results in an HBM effect. That is, in the presence of Wnt, G171F is able to activate the TCF-luciferase reporter construct. In fact, it may activate the reporter to a greater extent than either LRP5 or HBM. Furthermore, in the presence of Dkk1 and Wnt1, G171F is less susceptible than LRP5 to modulation by Dkk. These data exemplify that the G171F variant modulates Wnt signaling in a manner similar to HBM. In addition, this data confirms that HBM's valine residue at 171 is not the only modification at 171 that can result in an HBM effect. Together these data support an important role for LRP5 propeller 1 in modulating Wnt pathway activity; in responding to Dkk modulation; and, in the ability to generate an HBM effect.

Figure 25 shows the effects of the M282V mutation in propeller 1 of LRP5. M282 expression results in an HBM-effect. That is, in the presence of Wnt, M282 is able to activate the TCF-luciferase reporter construct. Furthermore, in the presence of Dkk1 and Wnt1, M282V is less susceptible than LRP5 to modulation by Dkk. These data show that the M282V variant modulates Wnt signaling in a manner similar to HBM. In addition, this data confirms that modifications of other residues in propeller 1 of LRP5 can result in an HBM effect.

These data support an "occupied space" model of the HBM mutation in propeller 1 and show that multiple mutations of propeller 1 are capable of generating an HBM effect; the original G171V HBM mutation is not unique in this ability. Moreover, various perturbations in propeller 1 can modulate Dkk activity.

These data illustrate the molecular mechanism of Dkk modulation of LRP signaling. Using the methods disclosed herein and in U.S. Application 60/290,071, generation of a comprehensive mutant panel will reveal residues in LRP that function in Dkk modulation of Wnt signaling. Such variants of LRP5 and LRP6 that modulate Dkk

activity and the residues which distinguish them from LRP5 and LRP6 are points for therapeutic intervention by small molecule compound, antibody, peptide aptamer, or other agents. Furthermore, models of each HBM-effect mutation/polymorphism may be used in rational drug design of an HBM mimetic agent.

These are only a few illustrative examples presented to better describe the present invention. Variants of LRP5 which have demonstrated HBM activity in assays include G171F, M282V, G171K, G171Q and A214V. Clearly, other variants may be contemplated within the scope of the present invention. Furthermore, wherever HBM is recited in the methods of the invention, it should be understood that any such alternative variant of LRP which demonstrates HBM biological activity is also encompassed by those claims.

17. <u>Screening Assays</u>

The two-hybrid system is extremely useful for studying protein:protein interactions. *See, e.g.,* Chien *et al., Proc. Natl Acad. Sci. USA* 88:9578-82 (1991); Fields *et al., Trends Genetics* 10:286-92 (1994); Harper *et al., Cell* 75:805-16 (1993); Vojtek *et al., Cell* 74:205-14 (1993); Luban *et al., Cell* 73:1067-78 (1993); Li *et al., FASEB J.* 7:957-63 (1993); Zang *et al., Nature* 364:308-13 (1993); Golemis *et al., Mol. Cell. Biol.* 12:3006-14 (1992); Sato *et al., Proc. Natl Acad. Sci. USA* 91:9238-42 (1994); Coghlan *et al., Science* 267:108-111 (1995); Kalpana *et al., Science* 266:2002-6 (1994); Helps *et al., FEBS Lett.* 340:93-8 (1994); Yeung *et al., Genes & Devel.* 8:2087-9 (1994); Durfee *et al., Genes & Devel.* 7:555-569 (1993); Paetkau *et al., Genes & Devel.* 8:2035-45; Spaargaren *et al.*, 1994 *Proc. Natl. Acad. Sci. USA* 91:12609-13 (1994); Ye *et al., Proc. Natl Acad. Sci. USA* 91:12629-33 (1994); and U.S. Patent Nos. 5,989,808; 6,251,602; and 6,284,519.

Variations of the system are available for screening yeast phagemid (*see, e.g.,* Harper, <u>Cellular Interactions and Development: A Practical Approach</u>, 153-179 (1993); Elledge *et al., Proc. Natl Acad. Sci. USA* 88:1731-5 (1991)) or plasmid (Bartel, 1993 and Bartel, *Cell* 14:920-4 (1993)); Finley *et al., Proc. Natl Acad. Sci. USA* 91:12980-4

(1994)) cDNA libraries to clone interacting proteins, as well as for studying known protein pairs.

The success of the two-hybrid system relies upon the fact that the DNA binding and polymerase activation domains of many transcription factors, such as GAL4, can be separated and then rejoined to restore functionality (Morin *et al.*, *Nuc. Acids Res.* 21:2157-63 (1993)). While these examples describe two-hybrid screens in the yeast system, it is understood that a two-hybrid screen may be conducted in other systems such as mammalian cell lines. The invention is therefore not limited to the use of a yeast two-hybrid system, but encompasses such alternative systems.

Yeast strains with integrated copies of various reporter gene cassettes, such as for example GAL.fwdarw.LacZ, GAL.fwdarw.HIS3 or GAL.fwdarw.URA3 (Bartel, in <u>Cellular Interactions and Development: A Practical Approach</u>, 153-179 (1993); Harper *et al.*, *Cell* 75:805-16 (1993); Fields *et al.*, *Trends Genetics* 10:286-92 (1994)) are co-transformed with two plasmids, each expressing a different fusion protein. One plasmid encodes a fusion between protein "X" and the DNA binding domain of, for example, the GAL4 yeast transcription activator (Brent *et al.*, *Cell* 43:729-36 (1985); Ma *et al.*, *Cell* 48:847-53 (1987); Keegan *et al.*, *Science* 231:699-704 (1986)), while the other plasmid encodes a fusion between protein "Y" and the RNA polymerase activation domain of GAL4 (Keegan *et al.*, 1986). The plasmids are transformed into a strain of the yeast that contains a reporter gene, such as lacZ, whose regulatory region contains GAL4 binding sites. If proteins X and Y interact, they reconstitute a functional GAL4 transcription activator protein by bringing the two GAL4 components into sufficient proximity to activate transcription. It is well understood that the role of bait and prey proteins may be alternatively switched and thus the embodiments of this invention contemplate and encompass both alternative arrangements.

Either hybrid protein alone must be unable to activate transcription of the reporter gene, the DNA-binding domain hybrid, because it does not provide an activation function, and the activation domain hybrid, because it cannot localize to the GAL4 binding sites. Interaction of the two test proteins reconstitutes the function of GAL4 and results in expression of the reporter gene. The reporter gene cassettes

consist of minimal promoters that contain the GAL4 DNA recognition site (Johnson *et al.*, *Mol. Cell. Biol.* 4:1440-8 (1984); Lorch *et al., J. Mol. Biol.* 186:821-824 (1984)) cloned 5' to their TATA box. Transcription activation is scored by measuring either the expression of β-galactosidase or the growth of the transformants on minimal medium lacking the specific nutrient that permits auxotrophic selection for the transcription product, *e.g.*, URA3 (uracil selection) or HIS3 (histidine selection). See, *e.g.*, Bartel, 1993; Durfee *et al.*, *Genes & Devel.* 7:555-569 (1993); Fields *et al.*, *Trends Genet.* 10:286-292 (1994); and U.S. Pat. No. 5,283,173.

Generally, these methods include two proteins to be tested for interaction which are expressed as hybrids in the nucleus of a yeast cell. One of the proteins is fused to the DNA-binding domain (DBD) of a transcription factor and the other is fused to a transcription activation domain (AD). If the proteins interact, they reconstitute a functional transcription factor that activates one or more reporter genes that contain binding sites for the DBD. Exemplary two-hybrid assays which have been used for Dkk-1 or Dkk-1/LRP5 are presented in the Examples below.

Additional methods of preparing two hybrid assay systems for Dkk-1 interactors would be evident to one of ordinary skill in the art. See for example, Finley *et al.*, "Two-Hybrid Analysis of Genetic Regulatory Networks," in The Yeast Two-Hybrid System (Paul L. Bartel et al., eds., Oxford, 1997); Meijia Yang, "Use of a Combinatorial Peptide Library in the Two-Hybrid Assay," in The Yeast Two-Hybrid System (Paul L. Bartel et al., eds., Oxford, 1997); Gietz *et al.*, "Identification of proteins that interact with a protein of interest: Applications of the yeast two-hybrid system," *Mol. & Cell. Biochem.* 172:67-9 (1997); K. H. Young, "Yeast Two-Hybrid: So Many Interactions,(in) so Little Time," *Biol. Reprod.* 58:302-311 (1998); R. Brent *et al.*, "Understanding Gene and Allele Function with Two-Hybrid Methods," *Annu. Rev. Genet.* 31:663-704 (1997). It will be appreciated that protein networks can be elucidated by performing sequential screens of activation domain-fusion libraries.

Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and utilize the compounds of the present invention and practice the claimed methods. The

following working examples therefore, specifically point out preferred embodiments of the present invention, and are not to be construed as limiting in any way the remainder of the disclosure.

EXAMPLES

The present invention is described by reference to the following Examples, which are offered by way of illustration and are not intended to limit the invention in any manner. Standard techniques well-known in the art or the techniques specifically described below were utilized.

For routine practice of the protocols referenced below, one of skill in the art is directed to the references cited in this application as well as the several Current Protocol guides, which are continuously updated, widely available and published by John Wiley and Sons, (New York). In the life sciences, Current Protocols publishes comprehensive manuals in Molecular Biology, Immunology, Human Genetics, Protein Science, Cytometry, Neuroscience, Pharmacology, Cell Biology, Toxicology, and Nucleic Acid Chemistry. Additional sources are known to one of skill in the art.

Example 1

Yeast Two Hybrid Screen Using LRP5 Ligand Binding Domain (LBD) Bait Sequences

In a screen against human osteoblast library (*i.e.*, HOB03C5, a custom Gibco generated Y2H compatible cDNA library from a human osteoblast cell line as described by Bodine and Komm, *Bone* 25:535-43 (1999)), an interaction with Dkk-1 was identified. The LRP5 ligand binding domain (LBD) baits used for this screen are depicted in Figures 2B and C. The basic protocol is as follows:

An overnight culture of the yeast strain containing the bait of interest is grown in 20 ml of appropriate selective medium containing 2% glucose at 30°C. The overnight culture is diluted by a 10 fold factor into YPDmedia supplemented with 40 mg/l of adenine, and grown for 4 hours at 30°C.

For each mating event, an aliquot of the frozen prey library is grown in 150 ml YAPD medium for 5 hours at 30°C.

Appropriate volumes calculated by measuring the OD600 of each culture are combined into a tube. The number of diploids to be screened is typically ten times the number of clones originally present in the prey library of interest. Assuming a mating efficiency of 20% minimum, fifty times (*i.e.*, ten times coverage multiplied by 20% mating efficiency) as many haploid cells containing the bait and as many cells containing the prey are used in any given mating event. The mixture is filtered over a 47 mm, 0.45 mm sterile Metricel filter membrane (Gelman).

Using sterile forceps, the filter is transferred onto a 100 mm^2 YAPD agar plate with the cell side up, removing all air bubbles underneath the filter. The plate is incubated overnight at room temperature.

The filter is transferred into a 50 ml Falcon tube using sterile forceps and 10 ml SD medium containing 2% glucose are added to resuspend the cells. The filter, once free of cells, is removed and the cell suspension is spun for 5 min. at 2,000 xg.

The cells are resuspended in 10 ml SD medium containing 2% glucose. An aliquot of 100 μl is set aside for titration.

The cells are plated onto large square plates containing appropriate selective media and incubated at 30°C for three to five days.

To calculate the mating efficiency and to determine the total number of diploid cells screened, the 100 μl aliquot set aside for titration is diluted and plated onto different selective media. The mating efficiency is calculated by dividing the number of diploids/ml by the lowest number of haploids/ml, either bait or prey, and multiplied by 100. For example, if 2 million diploids were obtained by mating 10 million of haploids containing a bait and 12 million of haploids containing a prey, then the mating efficiency is calculated by dividing 2 million by 10 million, which equals 0.2 and multiplied by 100 which equals 20%. Typical mating efficiencies under the above conditions are within about 20 to about 40%. The total number of diploids screened in a mating event is obtained by multiplying the number of diploids/ml by the total number of ml plated, typically about 10.

Isolation of colonies containing pairs of interacting proteins.

Yeast colonies from the interaction selection (large square) plates are picked with a sterile toothpick and patched onto plates containing the appropriate selective media and incubated at 30°C for two days.

To further ensure purity of the yeast, the plates are replicated onto another plate containing the same media and incubated at 30°C for another two days.

Yeast patches are scraped using a sterile toothpick and placed into a 96-well format plate containing 100 μl SD –L –W –H with 2% glucose liquid medium.

Half the volume of the plate is transferred to a 96-well plate containing 50 μl of 40% glycerol for storage. The other half is set aside for replication and galactosidase-activity assay (see below).

Cells are replicated onto a SD –L –W –H plate with 2% glucose plate to create a master plate, and incubated two days at 30°C. The master plate is replicated onto different selective media to score the strength of each interaction.

Cells are also replicated onto media selecting for the prey vector only for colony PCR and incubated two days at 30°C.

Galactosidase activity assay

Ten microliters from the 96-well plate (set aside from above) are transferred into another 96-well plate containing 100 μl SD and 2% glucose media. The cell density is measured at OD_{600} using a spectrophotometer, the OD_{600} is usually between 0.03 and 0.1. Fifty microliters of Galactosidase reaction mixture (Tropix) are added to microplates (Marsh) specifically designed for the luminometer (Hewlett Packard Lumicount). Fifty microliters of the diluted cells are then added and mixed by pipetting. The reaction is incubated sixty to one hundred twenty minutes at room temperature. Relative Light Units (RLUs) are read by the luminometer. Each plate contains a negative control, constituted by diploid yeast containing the bait of interest and an empty prey vector. To be scored as positive, the diploids tested have to have an RLU number at least twice as high as the negative control.

Example 2

Minimum interaction domain mapping

Further analysis of yeast two hybrid (Y2H) interacting proteins includes the dissection of protein motifs responsible for the interaction. Sequence alignment of multiple clones identified in the Y2H screens can help identify the smallest common region responsible for the interaction. In the absence of appropriate clones, deletion mapping of interacting domains is necessary.

PCR primers containing restriction sites suitable for cloning are designed to cover multiple sub-domains of the protein of interest (bait or prey). The methods involved in cloning, sequencing, yeast transformation, mating, and scoring of interactions are readily performed by one of ordinary skill in the art of molecular biology and genetic engineering.

Materials and Methods

Minimum interaction domain: primers were designed for PCR of the Dkk-1 clone isolated by screening a primary osteoblast cell strain (HOB03C5) library with pooled Zmax1/LRP5 ligand binding domain (LBD) baits: LBD1 (Leu969-Pro1376) and LBD4 (Arg1070-Pro1376). The primers, which are presented in 5' to 3' orientation, were as follows:

SEQ ID NO	Primer	Sequence
155	Forward 1	TTTTTTGTCGACCAATTCCAACGCTATCAAG
156	Forward 2	TTTTTTGTCGACCTGCGCTAGTCCCACCCGC
157	Forward 3	TTTTTTGTCGACCGTGTCTTCTGATCAAAATC
158	Forward 4	TTTTTTGTCGACCGGACAAGAAGGTTCTGTTTG
159	Reverse 1	TTTTTTGCGGCCGCTTATTTGGTGTGATACATTTTTG
160	Reverse 2	TTTTTTGCGGCCGCTTAGCAAGACAGACCTTCTCC
161	Reverse 3	TTTTTTGCGGCCGCTTAGTGTCTCTGACAAGTGTG

PCR was performed using PfuTurbo® polymerase (Stratagene). The PCR products were gel purified, digested with *SalI/ NotI* and ligated to pPC86 (Gibco/BRL) which had been linearized with *SalI/NotI*. Clones were recovered and sequenced to ascertain that the structure was as expected and that the Gal4 activation domain and Dkk-1 were in-frame. The ORF of Dkk-1 was Met1-His266, as in human Dkk-1 (GenBank Accession No. XM_005730).

The clones used were as follows: D5 (F1/R3: Asn34-His266), D4 (F1/R2: Asn34-Cys245), D3 (F1/R1: Asn34-Lys182), D9 (F2/R3: Cys97-His266), D12 (F3/R3, val139-His266), D14 (F4/R3: Gly183-His266), D8 (F2/R2: Cys97-Cys245), and D11 (F3/R2: Val139-Cys245). F1, F2, F3 and F4 refer respectively to Forward primers 1, 2, 3 and 4. R1, R2 and R3 refer respectively to reverse primers 1, 2 and 3.

These clones were transformed into yeast and mated with each of three yeast strains containing pDBleu (Gibco/BRL), pDBleuLBD1, and pDBleuLBD4. Positive interactions were detected by growth of the hybrids on appropriate selective media.

Results

Minimum interaction domain: Figure 6 shows that while growth was observed in diploids of D4, D5, D8, D9, and D12, no growth was observed in hybrids of D3, D11, and D12. Carboxy terminal (C-terminal) deletions indicated that while the C-terminal amino acids of Dkk-1 containing the potential N-glycosylation site (Arg246-His266) are not required for interaction with Zmax1/LRP5 LBD baits, the Cys2 domain, Gly183-Cys245, is required. N-terminal deletions also demonstrated that the region between the two cysteine domains, *i.e.* Val139 to Lys182, is also required. Two minimum interaction domain constructs were isolated: D12 (Val139-His266) and D8 (Cys97-Cys245). Similar constructs could be prepared for Dkk-1 interactors.

Example 3

Yeast-2 Hybrid screen for peptide aptamer sequences to Dkk-1

Peptide aptamer library construction

A peptide aptamer library, Tpep, was constructed, which provides a means to identify chimeric proteins that bind to a protein target (or bait) of interest using classic yeast two hybrid (Y2H) assays. The Tpep library is a combinatorial aptamer library composed of constrained random peptides, expressed within the context of the disulfide loop of *E. coli* thioredoxin (trxA), and as C-termini fusion to the *S. cerevisiae* Gal4 activation domain. The Tpep library was generated using a restriction enzyme modified recombinant Y2H prey vector, pPC86 (Gibco), which contains the trxA scaffold protein.

Generation of aptamer-encoding sequences

Aptamer-encoding sequences were produced as follows. DNA encoding random stretches of approximately sixteen amino acids surrounded by appropriate restriction sites were generated by semi-random oligonucleotide synthesis. The synthetic oligonucleotides were PCR-amplified, restriction digested, and cloned into the permissive sites within the trxA scaffold protein. The cloning strategy was to insert the random oligonucleotide sequence is in-frame with the scaffold protein coding sequence, resulting in expression of a scaffold protein-aptamer chimera. The scaffold protein is itself in-frame with the activation domain of Gal4, within the pPC86 vector that is appropriate for the aptamer to be expressed and functional in a regular Y2H assay. Additional methods of preparing aptamers would be apparent to the skilled artisan.

Generation of a permissive recombinant pPC86 vector containing the trxA coding sequence

First the *Rsr*II restriction site located within the Gal4 activation domain of pPC86 (Gibco) was eliminated by site-directed mutagenesis (Quickchange™ kit, Stratagene). The amino acid sequence of the Gal4 activation domain was unchanged by this modification. The strength of different control interactions was verified to be unchanged by the modification.

Second, the *E. coli* trxA coding sequence was cloned into the *Sal*I and *Not*I sites of the *Rsr*II-modified pPC86. *Eco*RI and *Spe*I sites were then introduced within the trxA

*Rsr*II site. The oligonucleotides encoding the peptide aptamers were cloned into the *Eco*RI and *Spe*I sites of the resulting vector.

Example 4

Yeast-2 Hybrid screen for Dkk-1 interacting proteins

A Dkk-1 bait sequence was utilized in a yeast two hybrid screen to identify Dkk-1 interacting proteins. The procedure for the Y2H was carried out similarly to that employed in Example 1, except that the Dkk-1 bait from Figure 2C was used instead of LRP baits. The screen was performed using Hela and fetal brain libraries (Invitrogen Corporation, Carlsbad, CA). Multiple libraries were used to identify additional Dkk-1 interacting proteins and to confirm interactions found in other libraries.

The list of Dkk-1 interacting proteins uncovered in these Y2H screens are listed in Figure 5.

The interacting proteins identified in the Dkk-1 bait screen can be used in other Y2H screens with LRP baits and other Dkk-1 interacting proteins to determine more complex interactions which may modulate Dkk-1/LRP interactions and/or Wnt signaling.

Example 5

Generation of antibodies

In each of the following antibody-generating examples, the synthesis of these linear peptides is followed by injection into two New Zealand Rabbits. Subsequent boosts and bleeds are taken according to a standard ten-week protocol. The end-user receives back 5 mgs of peptide, aliquots of pre-bleeds, roughly 80 ml of crude sera from each of the two rabbits and, and ELISA titration data is obtained.

Generation of LRP5 Polymorphism-specific antibodies

Antibodies were generated to the following peptides to obtain antibodies which distinguish the HBM polymorphism versus wild-type LRP5/Zmax: MYWTDWVETPRIE

(SEQ ID NO:123) (mutant peptide) and MYWTDWGETPRIE (SEQ ID NO:124) (wild-type peptide for negative selection). Immunofluorescence data confirmed that the antibody, after affinity purification, is specific for HBM and does not recognize LRP5 (Figure 17).

Generation of LRP5 Monospecific antibodies

LRP5 monospecific polyclonal antibodies were generated to the following amino acid sequences of LRP5: Peptide 1 (a.a. 265-277) - KRTGGKRKEILSA (SEQ ID NO:125), Peptide 2 (a.a. 1178-1194) - ERVEKTTGDKRTRIQGR (SEQ ID NO:126), and Peptide 3 (a.a. 1352-1375) - KQQCDSFPDCIDGSDE (SEQ ID NO:127). Immunofluorescence confirmed that the antibody generated detects LRP5.

Generation of Dkk-1 monospecific polyclonal antibodies

Dkk-1 monospecific polyclonal antibodies were generated to the following amino acid sequences of Dkk-1: Peptide 1 (a.a. 71-85) - GNKYQTIDNYQPYPC (SEQ ID NO:118), Peptide 2 (a.a. 165-186) - LDGYSRRTTLSSKMYHTKGQEG (SEQ ID NO:119),.Peptide 3 (a.a. 246-266) - RIQKDHHQASNSSRLHTCQRH (SEQ ID NO:120), Peptide 4 (a.a. 147-161) - RGEIEETITESFGND (SEQ ID NO:121), and Peptide 5 (232-250) - EIFQRCYCGEGLSCRIQKD (SEQ ID NO:122) of human Dkk-1. Figure 26 shows the location of the various peptides selected, their relationship to the Dkk-1 amino acid sequence and polyclonal antibodies generated.

Western blots demonstrated that the antibodies generated against peptides 2 (Antibody #5521) (Figure 27) and 4 (Antibody #74397) (Figure 28) are specific toward Dkk-1. Figure 27 shows Western blots using 500 μl of conditioned medium (CM) from non-transfected 293 cells or from 293 cells transfected with Dkk1-V5 that were immunoprecipitated by anti-V5 antibody. Bead elutes were separated by non-reducing SDS-PAGE (lanes #4, 5 of Figure 27). 20 μl of conditioned medium from both samples (lanes #2, 3 of Figure 27) and from Dkk1-AP transfected 293 cells (lane #6 of Figure 27) were additionally separated on the gel. The Western was performed using

antibodies Anti-V5/AP (1:10,000) and Ab#5521 (10 μg/ml). Ab#5521 detected Dkk1-V5 and Dkk1-AP from conditioned medium.

Figure 28 shows Western blot results using Ab#74397. Anti-V5/AP was tested at a 1:4000 dilution and Ab#74397 was tested at a 1:500 dilution. Ab#74397 was able to detect Dkk1-V5 in both conditioned medium and immunoprecipitated conditioned medium.

The results obtained with antibodies #5521 and #74397 are summarized in the following table:

Rabbit No.	Peptide Position	Peptide Sequence	Purified (Y/N)	Western	Immuno-precipitation	Location
5521	165-186	LDGYSR RTTLSSK MYHTKG QEG	Y (Protein G purified)	Y	N/A	Between Cy1 and Cys2 domain
74397	147-161	RGEIEETI TESFGN D	N	Y	N/A	Between Cy1 and Cys2 domain

Example 6

Effects of exogenous Dkk-1 on Wnt-mediated signaling in the Xenopus embryo assay

Xenopus embryos are an informative and well-established *in vivo* assay system to evaluate the modulation of Wnt signaling (McMahon *et al.*, *Cell* 58: 1075-84 (1989); Smith and Harland, 1991; reviewed in Wodarz and Nusse 1998) .

Modification of the Wnt signaling pathway can be visualized by examining the embryos for a dorsalization phenotype (duplicated body axis) after RNA injection into the ventral blastomere at the 4- or 8-cell stage. On the molecular level, phenotypes can be analyzed by looking for expression of various marker genes in stage 10.5 embryos. Such markers would include general endoderm, mesoderm, and ectoderm markers as well as a variety of tissue-specific transcripts.

Analysis can be done by RT-PCR/TaqMan® and can be done on whole embryo tissue or in a more restricted fashion (microdissection). Because this system is very flexible and rapid, by injecting combinations of transcripts, such as HBM and different Wnts or Wnt antagonists, the mechanism of HBM in the Wnt pathwaycan thereby be dissected. Furthermore, investigations are conducted to determine whether Zmax/LRP5 and HBM differentially modulate Wnt signaling either alone, or in combination with other components. Previous studies have demonstrated that LRP6 alone or LRP5 + Wnt5a were able to induce axis duplication (dorsalization) in this system (Tamai *et al.*, *Nature* 407: 530-35 (2000)).

Constructs for Xenopus Expression (Vector pCS2⁺)

Constructs were prepared using the vector pCS2⁺. DNA inserts were subcloned in the sense orientation with respect to the vector SP6 promoter. The pCS2⁺ vector contains an SV40 virus polyadenylation signal and T3 promoter sequence (for generation of antisense mRNA) downstream of the insert.

Full length Zmax/LRP5 and HBM ORF cDNA: Insert cDNA was isolated from the full length cDNA retrovirus constructs (with optimized Kozak sequences) by *Bgl*II-*Eco*RI digestion and subcloned into the *Bam*HI-*Eco*RI sites of the pCS2⁺ vector.

Full length XWnt8: This cDNA was PCR amplified from a Xenopus embryo cDNA library using oligos 114484 (SEQ ID NO:162) (5'-CAGTGAATTCACCATGCAAAACACCACTTTGTTC-3') and 114487 (SEQ ID NO:163) (5'-CAGTTGCGGCCGCTCATCTCCGGTGGCCTCTG-3'). The oligos were designed to amplify the ORF with a consensus Kozak sequence at the 5' end as determined from GenBank #X57234. PCR was carried out using the following conditions: 96°C, 45 sec.; 63°C, 45 sec.; 72°C, 2 min. for 30 cycles. The resulting PCR product was purified, subcloned into pCRII-TOPO (Invitrogen Corp.), sequence verified, and digested with *Bam*HI/*Xho*I. This insert was subcloned into the vector at the *Bam*HI-*Xho*I sites.

Full length Wnt5a: A murine Wnt5a cDNA clone was purchased from Upstate Biotechnology (Lake Placid, NY) and subcloned into the *Eco*RI site of the vector. Sequencing confirmed insert orientation.

Full length human Dkk-1: A human cDNA with GenBank accession number AF127563 was available in the public database. Using this sequence, PCR primers were designed to amplify the open reading frame with a consensus Kozak sequence immediately upstream of the initiating ATG. Oligos 117162 (SEQ ID NO:164) (5'-CAATAGTCGACGAATTCACCATGGCTCTGGGCGCAGCGG-3') and 117163 (SEQ ID NO:165) (5'-GTATTGCGGCCGCTCTAGATTAGTGTCTCTGACAAGTGTGAA-3') were used to screen a human uterus cDNA library by PCR. The resulting PCR product was purified, subcloned into pCRII-TOPO (Invitrogen Corp.), sequence verified, and digested with *EcoRI/XhoI*. This insert was subcloned into the pCS2$^+$ vector at the *EcoRI-XhoI* sites.

Full length human Dkk-2: A full length cDNA encoding human Dkk-2 was isolated to investigate the specificity of the Zmax/LRP5/HBM interaction with the Dkk family of molecules. Dkk-1 was identified in yeast as a potential binding partner of Zmax/LRP5/HBM. Dkk-1 has also been shown in the literature to be an antagonist of the Wnt signaling pathway, while Dkk-2 is not (Krupnik *et al.*, 1999). The Dkk-2 full length cDNA serves as a tool to discriminate the specificity and biological significance of Zmax/LRP5/HBM interactions with the Dkk family (*e.g.*, Dkk-1, Dkk-2, Dkk-3, Dkk-4, Soggy, their homologs and variant, etc.). A human cDNA sequence for Dkk-2 (GenBank Accession No. NM_014421) was available in the public database. Using this sequence, PCR primers were designed to amplify the open reading frame with a consensus Kozak sequence immediately upstream of the initiating ATG. Oligos 51409 (SEQ ID NO:166) (5'- CTAACGGATCCACCATGGCCGCGTTGATGCGG-3') and 51411 (SEQ ID NO:167) (5'-GATTCGAATTCTCAAATTTTCTGACACACATGG-3') were used to screen human embryo and brain cDNA libraries by PCR. The resulting PCR product was purified, subcloned into pCRII-TOPO, sequence verified, and digested with *Bam*HI/*EcoR*I. This insert was subcloned into the pCS2$^+$ vector at the *Bam*HI-*EcoR*I sites.

Full length LRP6 was isolated from the pED6dpc4 vector by *XhoI-XbaI* digestion. The full length cDNA was reassembled into the *XhoI-XbaI* sites of pCS2$^+$. Insert orientation was confirmed by DNA sequencing.

mRNA Synthesis and Microinjection Protocol

mRNA for microinjection into *Xenopus* embryos is generated by *in vitro* transcription using the cDNA constructs in the pCS2^{+} vector described above as template. RNA is synthesized using the Ambion mMessage mMachine high yield capped RNA transcription kit (Cat. #1340) following the manufacturer's specifications for the Sp6 polymerase reactions. RNA products were brought up to a final volume of 50 µl in sterile, glass-distilled water and purified over Quick Spin Columns for Radiolabelled RNA Purification G50-Sephadex (Roche, Cat. #1274015) following the manufacturer's specifications. The resulting eluate was finally extracted with phenol:chloroform:isoamyl alcohol and isopropanol precipitated using standard protocols (Sambrook *et al.*, 1989). Final RNA volumes were approximately 50 µl. RNA concentration was determined by absorbance values at 260 nm and 280 nm. RNA integrity was visualized by ethidium bromide staining of denaturing (formaldehyde) agarose gel electrophoresis (Sambrook *et al.*, 1989). Various amounts of RNA (2 pg to 1 ng) are injected into the ventral blastomere of the 4- or 8-cell *Xenopus* embryo. These protocols are described in Moon *et al.*, *Technique-J. of Methods in Cell and Mol. Biol.* 1: 76-89 (1989), and Peng, *Meth. Cell. Biol.* 36: 657-62 (1991).

Screening for Duplicated Body Axis

In vitro transcribed RNA is purified and injected into a ventral blasomere of the 4- or 8-cell *Xenopus* embryo (approx. 2 hours post-fertilization). At stage 10.5 (approx. 11 hours post-fertilization), the injected embryos are cultured for a total of 72 hours and then screened for the presence of a duplicated body axis (dorsalization) (Figure 7). Using XWnt8-injected (2-10 pg) as a positive control (Christian et al. (1991)) and water-injected or non-injected embryos as negative controls, we replicated the published observation that Zmax(LRP5) + Wnt5a (500 and 20 pg, respectively) could induce axis duplication. Wnt5a (20 pg) alone could not induce axis duplication (as previously reported by Moon *et al.* (1993)). We have also injected GFP RNA (100-770 pg) as a negative control to show that the amount of RNA injected is not perturbing embryo development (not shown). Strikingly, HBM + Wnt5a (500 and 20 pg, respectively)

yielded an approximately 3.5 fold more robust response of the phenotype ($p=0.043$ by Fisher's exact test) compared to Zmax(LRP5) + Wnt5a, suggesting that the HBM mutation is activating the Wnt pathway (Figures 8 and 9). The HBM/Wnt5a embryos also appear to be more "anteriorized" than the Zmax(LRP5)/Wnt5a embryos, again suggestive of a gain-of-function mutation.

The role of Dkk-1 as a modulator of Zmax/LRP5- and HBM-mediated Wnt signaling was investigated. Literature reports have previously characterized *Xenopus* and murine Dkk-1 as antagonists of the canonical Wnt pathway in the *Xenopus* system (Glinka *et al.*, *Nature* 391:357-362 (1998)). Using the human Dkk-1 construct, a dose-response assay was performed to confirm that our construct was functional and to identify the optimal amount of RNA for microinjection. Using 250 pg/embryo of hDkk-1 RNA, over 90% ($p<0.001$) of the embryos were observed to display enlarged anterior structures (big heads) as anticipated from the published reports (Figure 10).

The mechanism of hDkk-1 modulation of Wnt signaling in the presence of Zmax/LRP5 or HBM was also investigated. Without any hDkk-1 present, it was confirmed that HBM + Wnt5a was a more potent activator of Wnt signaling than Zmax/LRP5 + Wnt5a ($p<0.05$). Interestingly, in the presence of hDkk-1 (250 pg), Zmax/LRP5-mediated Wnt signaling was repressed ($p<0.05$) but hDkk-1 was unable to repress HBM-mediated Wnt signaling ($p<0.01$) (Figure 11). The specificity of this observation can be further addressed by investigating other members of the Dkk family, other Wnt genes, LRP6, additional Zmax/LRP5 mutants, and the peptide aptamers.

Example 7

Effects of exogenous Dkk and LRP5 on Wnt signaling in the TCF-luciferase Assay

Wnt activity can be antagonized by many proteins including secreted Frizzled related proteins (SFRPs), Cerberus, Wnt Inhibitory Factor-1 and Dkk-1 (Krupnik *et al.*, 1999). The Dkk family of proteins consists of Dkk-1-4 and Soggy, a Dkk-3-like protein. Dkk-1 and Dkk-4 have been shown to antagonize Wnt mediated *Xenopus* embryo development, whereas Dkk-2, Dkk-3, and Soggy do not. Unlike many of these proteins

that antagonize Wnt activity by directly interacting with Wnt proteins, Dkk-1 acts by binding to two recently identified Wnt coreceptors, LRP5 and LRP6. (Mao *et al.*, 2001; Bafico *et al.*, 2001). The details of this interaction have been examined by the present inventors and Mao et al. using deletion constructs of LRP6, which demonstrated that EGF repeats 3 and 4 are important for Dkk-1 interaction. Accordingly, the activity of two Dkk proteins, Dkk-1 and Dkk-2, were investigated with various Wnt members, LRP5, LRP6, and the mutant form of LRP5, designated HBM. The present invention explores whether there is any functional difference between LRP5 and HBM with regard to Dkk action on Wnt mediated signaling. Various reagents were developed, including Dkk-1 peptides, constrained LRP5 peptide aptamers, constrained Dkk-1 peptide aptamers and polyclonal antibodies to Dkk-1 (in Example 5 above) to identify factors that mimic HBM mediated Wnt signaling.

Methods

Various LRP5 constrained peptides were developed. Specifically, four peptides that interact with the LBD of LRP5 (Figure 4,constructs OST259-262 in Figure 12) and three peptides that interact with the cytoplasmic domain of LRP5 (constructs OST266-OST268 in Figure 12). In addition two Dkk-1 peptides were developed: constructs OST264 and OST265 in Figure 12, corresponding to Dkk-1 amino acids 139-266 and 96-245, containing the smallest region of Dkk-1 that interacts with LRP5 (Figure 6). The cDNA clones encoding the LRP5 LBD interacting peptides and the Dkk-1 peptides were subcloned into pcDNA3.1 with the addition of a Kozak and signal sequence to target the peptide for secretion. The constructs encoding the three peptides interacting with the cytoplasmic domain of LRP5 were also subcloned into pcDNA3.1. However, these latter constructs do not contain a signal sequence.

HOB-03-CE6 osteoblastic cells developed by Wyeth Ayerst (Philadelphia, PA) were seeded into 24-well plates at 150,000 cells per well in 1 ml of the growth media (D-MEM/F12 phenol red-free) containing 10% (v/v) heat-inactivated FBS, 1X penicillin streptomycin, and 1X Glutamax-1, and incubated overnight at 34°C. The following day, the cells were transfected using Lipofectamine 2000® (as described by the

manufacturer, Invitrogen) in OptiMEM (Invitrogen) with 0.35 μg /well of LRP5, HBM, or control plasmid DNA (empty vector pcDNA3.1) and either Wnt1 or Wnt3a plasmid DNA. Similar experiments were performed with LRP6 plasmid DNA (0.35μg/well) or a control pEDdpc4 empty vector. Furthermore, each of these groups were then divided into three groups, those receiving 0.35 μg/well Dkk-1, Dkk-2, or pcDNA3.1 control DNA. All wells were transfected with 0.025 μg/well of CMV beta-galactosidase plasmid DNA and 0.35 μg/well 16X TCF(AS)-luciferase reporter DNA (developed by Ramesh Bhat, Wyeth-Ayerst (Philadelphia, PA)). After 4 hours of incubation, the cells were rinsed and 1 ml of fresh growth media was added to each well. The cells were cultured overnight at 34°C, followed by a wash and a change of media. Cells were cultured for an additional 18-24 hours at 37°C. Cells were then lysed with 50 μl/well of 1X lysis buffer. The extracts were assayed for beta-galactosidase activity (Galacto Reaction Buffer Diluent & Light Emission Accelerator, Tropix) using 5 μl extract + 50 μl beta-galactosidase diluent and luciferase activity (Luciferase Assay Reagent, Promega) using 20 μl extract.

U2OS human osteosarcoma cells were also utilized. U2OS cells (ATCC) were seeded into 96-well plates at 30,000 cells per well in 200ul of the growth media (McCoy's 5A) containing 10% (v/v) heat-inactivated FBS, 1X penicillin streptomycin, and 1X Glutamax-1, and incubated overnight at 37°C. The following day, the nmedia was replaced with OptiMEM (Invitroge) and cells were transfected using Lipofectamine 2000® (as described by the manufacturer, Invitrogen) with 0.005µg/well of LRP5, HBM, LRP6 or contol plasmid DNA (empty vector pcDNA3.1) and either Wnt1 (.0025ug/well) or Wnt3a (.0025ug/well) plasmid DNA. In addition, the 16x-(AS) TCF-TK-firefly-luciferase (Ramesh Bhat, WHRI, Wyeth) and control TK-renilla luciferase (Promega Corp.) were co-transfected at 0.3ug/well and 0.06ug/well respectively in all experiments. Futhermore, each of these groups was then divided into different groups, those receiving 0.05ug/well Dkk-1, Dkk-2, Dkk3, Dkk1-Alkaline Phosphatase (AP), mutant Dkk-1 (C220A), Soggy or pcDNA3.1 control DNA. In other experiments, cells were co-transfected with 0.005 µg/well of LRP5, 0.0025ug/well of Wnt1 or Wnt3a (using 0.0025 µg/well of a control pcDNA3.1) with LRP5-interacting aptamers (0.05ug/well).

Cells were cultured for an additional 18-20 hours at 37°C. Culture medium was removed. Cells were cultured for an additional 18-20 hours at 37°C. Culture medium was removed. Cells were then lysed with 100 μl/well of 1X Passive Lysis Buffer (PLB) of Dual Luciferase Reagent kit (DLR-kit-Promega Corp.) 20 μl of the lysates were combined with LARII reagent of DLR-kit and assayed for TCF-firefly luciferase signal in Top Count (Packard) instrument. After measuring the Firefly readings, 100 μl of the "Stop and Glo" reagent of DLR kit that contains a quencher and a substrate for renilla luciferase was added into each well. Immediately the renilla luciferase reading was measured using the Top Count (Packard) Instrument. The ratios of the TCF-firefly luciferase to control renilla readings were calculated for each well and the mean ratio of triplicate or more wells was expressed in all data.

Results

The results of these experiments demonstrate that Dkk-1, in the presence of Wnt1 and LRP5, significantly antagonized TCF-luciferase activity (Figure 14). In marked contrast, Dkk-1 had no effect on HBM/Wnt1 mediated TCF-luciferase activity (Figure 14). In similar experiments, Dkk-1 was also able to antagonize LRP5/Wnt3a but not HBM/Wnt3a mediated TCF-luciferase activity (Figure 15). These results indicate that the HBM mutation renders Dkk-1 inactive as an antagonist of Wnt1 and Wnt3a signaling in HOB03CE6 osteoblastic cells. In other experiments with Wnt1, Dkk-1 had no effect on LRP5 or HBM mediated TCF-luciferase activity (Figure 14). In contrast, with either LRP5 or HBM in the presence of Wnt3a, Dkk-2 was able to antagonize the TCF-luciferase activity (Figure 15). These latter results indicate that the HBM mutation has no effect on Dkk-2 action in the presence of Wnt3a. Experiments were also performed using the closely related LRP6 cDNA in HOB-03-CE6 cells. In these experiments, LRP6/Wnt1 and LRP6/Wnt3a mediated TCF-luciferase were regulated in the same manner as LRP5. Specifically, Dkk-1 antagonized LRP6/Wnt1 mediated TCF-luciferase activity, whereas Dkk-2 had no effect (Figure 14). However, similar to the action of Dkk-2 with LRP5/Wnt3a, Dkk-2 was able to antagonize LRP6/Wnt3a mediated TCF-luciferase activity (Figure 15).

The results in the U2OS cells show a robust effect of the OST262 LRP5 peptide aptamer activation of Wnt signaling in the presence of Wnt3a (Figure 16). These functional results are confirmed by the results shown below in Example 11 using LRP5 peptide aptamers in the Xenopus assay. Such results affirmatively demonstrate that the effects of small molecules on LRP5/LRP6/HBM signaling can be detected using the TCF-luciferase assay.

These data demonstrate that there is a functional difference between LRP5 and HBM regarding the ability of Dkk-1 to antagonize Wnt1 and Wnt3a signaling. These data and previous data showing that Dkk-1 directly interacts with LRP5 suggests that the inability of Dkk-1 to antagonize HBM/Wnt signaling may in part contribute to the HBM phenotype. These experiments further demonstrate the ability to test various molecules (e.g., small molecules, aptamers, peptides, antibodies, LRP5 interacting proteins or Dkk-1 interacting proteins, and the like) for a LRP5 ligand that mimics HBM mediated Wnt signaling or factors that block Dkk-1 interaction with LRP5.

Example 8

Yeast-2 Hybrid Interaction Trap

Small molecule inhibitors (or partial inhibitors) of the Dkk-LRP interaction may be an excellent osteogenic therapeutic. One way to investigate this important protein-protein interaction is using Y2H techniques substantially as described above and as is well known in the art. Regions of LRP5, such as LRP5 LBD, have been found to functionally interact with Dkk. This interaction is quantitated using a reporter element known in the art, e.g., LacZ or luciferase, which is only activated when bait and prey interact. The Y2H assay is used to screen for compounds which modulate the LRP-Dkk interaction. Such a modulation would be visualized by a reduction in reporter element activation signifying a weaker or disrupted interaction, or by an enhancement of the reporter element activation signifying a stronger interaction. Thus, the Y2H assay can be used as a high-throughput screening technique to identify compounds which disrupt or enhance Dkk interaction with LRP5/LRP6/HBM, which may serve as potential therapeutics.

For example, the Interaction Trap methodology can be used as follows. The LRP5 LBD, for example, was fused with LexA and Dkk-1 was fused with either Gal4-AD or B42. With the LRP5LBD-LexA bait and the Gal4AD-Dkk prey, over a 20-fold activation of a lacZ reporter (under the control of a single LexA operator) was detected over the background. Using a Dkk-1 mutant (C220A) that is unable to bind to LRP, the interaction was reduced in yeast, showing the specificity of this interaction and system (Figure 18). As a result, small molecules may be identified that modulate this interaction between LRP and Dkk.

Example 9

Cell-Based Functional High-Throughput Assay

To develop a high throughput assay, the TCF-luciferase assay described in Example 7 was modified utilizing low level expression of endogenous LRP5/6 in U2OS and HEK293 cells. However, HOB-03-CE6 cells and any other cells which show a differential response to Dkk depending on whether LRP5, LRP6 or HBM are expressed. Using U2OS (human osteosarcoma) and HEK293 (ATCC) cells, the TCF-luciferase and tk-Renilla reporter element constructs were co-transfected along with Wnt3a/1 and Dkk. Wnt3a alone, by using endogenous LRP5/6, was able to stimulate TCF reporter gene activation. When Dkk, is co-transfected with Wnt3a/Wnt 1 and reporters (TCF-luci and tk-Renilla), Dkk represses reporter element activity. In addition, the TCF-luci signal is activated by Wnt3a/Wnt1 can be repressed by the addition of Dkk-enriched conditioned media to the cells containing Wnt3a/Wnt1 and reporters. The assay is further validated by the lack of TCF-reporter inhibition by a point mutant construct (C220A) of Dkk1.

The Dkk-mediated repression of the reporter is dependent upon the concentration of transfected Dkk cDNA or on the amount of Dkk-conditioned media added. In addition, the Dkk-mediated reporter suppression can be altered by the co-transfection of LRP5, LRP6, and HBM cDNAs in the U2OS or HEK293 cells. In general, U2OS cells show greater sensitivity to Dkk-mediated reporter suppression than that in HEK-293 cells. In U2OS cells, the transfection of LRP5/LRP6/HBM cDNA leads

to moderate activation of TCF-luci in the absence of Wnt3a/Wnt1 transfection. This activation presumably utilizes the endogenous Wnts present in U2OS cells. Under this condition, Dkk1 can repress TCF-luci and shows a differential signal between LRP5 and HBM. By co-transfecting Wnt3a/Wnt1, there is a generalized increase in the TCF-luci signal in the assay. Further, one can detect Dkk-mediated differential repression of the reporter due to LRP5 and HBM cDNA expression as well as between LRP5 and LRP6 cDNA. The repression is maximal with LRP6, moderate with LRP5, and least with HBM cDNA expression. In addition, the assay can detect the functional impact of the LRP5 interacting peptide aptamers (Figure 4), Dkk1 interacting aptamers and binding domains of Dkk-1 (Figure 6; OST264 and OST265 of Figures 12 and 13).

Using this system with a suppressed Wnt-TCF signal due to the presence of both Dkk and Wnt3a, one can screen for compounds that could alter Dkk modulation of Wnt signaling, by looking for compounds that activate or the TCF-luciferase reporter, and thereby relieve the Dkk-mediated repression of the Wnt pathway. Such compounds identified may potentially serve as HBM-mimetics and be useful, for example, as osteogenic therapeutics. Data generated from this high throughput screen are demonstrated in Figures 19-21. Figure 19 shows that Dkk1 represses Wnt3a-mediated signaling in U2OS bone cells. Figure 20 demonstrates the functional differences between LRP5, LRP6, and HBM. Dkk-1 represses LRP6 and LRP5 but has little or no effect on HBM-generated Wnt1 signaling in U2OS cells. Figure 21 demonstrates the differential effects of various Dkk family members and modified Dkks, including Dkk-1, a mutated Dkk-1 (C220A), Dkk-1-AP (modified with alkaline phosphatase), Dkk-3, and Soggy.

Example 10

DKK/LRP5/6/HBM ELISA Assay

A further method to investigate Dkk binding to LRP is via ELISA assay. Two possible permutations of this assay are exemplified. LRP5 is immobilized to a solid surface, such as a tissue culture plate well. One skilled in the art will recognize that other supports such as a nylon or nitrocellulose membrane, a silicon chip, a glass slide,

beads, etc. can be utilized. In this example, the form of LRP5 used is actually a fusion protein where the extracellular domain of LRP5 is fused to the Fc portion of human IgG. The LRP5-Fc fusion protein is produced in CHO cell extracts from stable cell lines. The LRP5-Fc fusion protein is immobilized on the solid surface via anti-human Fc antibody or by Protein-A or Protein G-coated plates, for example. The plate is then washed to remove any non-bound protein. Conditioned media containing secreted Dkk protein or secreted Dkk-epitope tagged protein (or purified Dkk or purified Dkk-epitope tagged protein) is incubated in the wells and binding of Dkk to LRP is investigated using antibodies to either Dkk or to an epitope tag. Dkk-V5 epitope tagged protein would be detected using an alkaline phosphatase tagged anti-V5 antibody.

Alternatively, the Dkk protein could be directly fused to a detection marker, such as alkaline phosphatase. Here the detection of the Dkk-LRP interaction can be directly investigated without subsequent antibody-based experiments. The bound Dkk is detected in an alkaline phosphatase assay. If the Dkk-alkaline phosphatase fusion protein is bound to the immobilized LRP5, alkaline phosphatase activity would be detected in a colorimetric readout. As a result, one can assay the ability of small molecule compounds to alter the binding of Dkk to LRP using this system. Compounds, when added with Dkk (or epitope-tagged Dkk) to each well of the plate, can be scored for their ability to modulate the interaction between Dkk and LRP based on the signal intensity of bound Dkk present in the well after a suitable incubation time and washing. The assay can be calibrated by doing cold competition experiments with unlabeled Dkk or with a second type of epitope-tagged Dkk. Any small molecule that is able to modulate the Dkk-LRP interaction may be a suitable therapeutic candidate, more preferably an osteogenic therapeutic candidate.

Example 11

Functional Evaluation of Peptide Aptamers in Xenopus

The constrained peptide aptamers constructs OST258-263 (where 258 contains the signal sequence by itself and 263 contains an irrelevant constrained peptide) (Figures 12 and 13) were used to generate RNA substantially as described in Example

7, except the vector was linearized by restriction endonuclease digestion and RNA was generated using T7 RNA polymerase.

Aptamer RNA was injected at 250 pg per blastomere using the protocol of Example 7. Wnt signaling was activated, as visualized by embryo dorsalization (duplicated body axis) with aptamers 261 and, more strongly, 262. The results of this assay are shown in Figures 22 and 23. These results suggest that aptamers 261 and 262 are able to activate Wnt signaling possibly by binding to the LBD of LRP, thereby preventing the modulation of LRP-mediated signaling by Dkk.

The aptamers of the present invention can serve as HBM-mimetics. In the Xenopus system they are able to induce Wnt signaling all by themselves. They may also serve as tools for rational drug design by enhancing the understanding of how peptides are able to interact with LRP and modulate Wnt signaling at the specific amino acid level. Thus, one would be able to design small molecules to mimic their effects as therapeutics. In addition, the aptamers identified as positives in this assay may be used as therapeutic molecules themselves.

Example 12

Homogenous Assay

An excellent method to investigate perturbations in protein-protein interactions is via Fluorescence Resonance Energy Transfer (FRET). FRET is a quantum mechanical process where a fluorescent molecule, the donor, transfers energy to an acceptor chromophore molecule which is in close proximity. This system has been successfully used in the literature to characterize the intermolecular interactions between LRP5 and Axin (Mao et al., *Molec. Cell Biol.* 7:801-809). There are many different fluorescent tags available for such studies and there are several ways to fluorescently tag the proteins of interest. For example, CFP (cyan fluorescent protein) and YFP (yellow fluorescent protein) can be used as donor and acceptor, respecively. Fusion proteins, with a donor and an acceptor, can be engineered, expressed, and purified.

For instance, purified LRP protein, or portions or domains thereof, fused to CFP and purified Dkk protein, or portions or domains thereof that interact with Dkk or LRP

respectively, fused to YFP can be generated and purified using standard approaches. If LRP-CFP and Dkk-YFP are in close proximity, the transfer of energy from CFP to YFP will result in a reduction of CFP emission and an increase in YFP emission. Energy is supplied with an excitation wavelength of 450 nm and the energy transfer is recorded at emission wavelengths of 480 nm and 570 nm. The ratio of YFP emission to CFP emission provides a guage for changes in the interaction between LRP and Dkk. This system is amenable for screening small molecule compounds that may alter the Dkk-LRP protein-protein interaction. Compounds that disrupt the interaction would be identified by a decrease in the ratio of YFP emission to CFP emission. Such compounds that modulate the LRP-Dkk interaction would then be considered candidate HBM mimetic molecules. Further characterization of the compounds can be done using the TCF-luciferase or Xenopus embryo assays to elucidate the effects of the compounds on Wnt signaling.

While the above example describes a cell-fee, solution-phase assay using purified components, a similar cell-based assay could also be performed. For example, LRP-CFP fusion protein can be expressed in cells. The Dkk-YFP fusion protein then could be added to the cells either as purified protein or as conditioned media. The interaction of LRP and Dkk is then monitored as described above.

All references cited herein are hereby incorporated by reference in their entirety for all purposes. The following applications are also incorporated by reference in their entirety herein for all purposes: U.S. Application No. 60/290,071, filed May 11, 2001; U.S. Application No. 09/544,398, filed on April 5, 2000; U.S. Application No. 09/543,771, filed April 5, 2000; 09/578,900; U.S. Application No. 09/229,319, filed January 13, 1999; U.S. Provisional Application 60/071,449, filed January 13, 1998; and International Application PCT/US00/16951, filed June 21, 2000; International PCT Application entitled "HBM Variants That Modulate Bone Mass and Lipid Levels," filed May 13, 2002; and International PCT Application entitled "Transgenic Animal Model of Bone Mass Modulation," filed May 13, 2002. Additionally, this application claims priority to U.S. provisional applications 60/291,311, filed May 17, 2001; 60/353,058, filed

February 1, 2002; and 60/361,293, filed March 4, 2002; the texts of which are herein incorporated by reference in their entirety for all purposes.

CLAIMS

We claim:

1. A method of regulating LRP5, LRP6, or HBM activity in a subject comprising administering a composition which modulates a Dkk activity in an amount effective to regulate LRP5, LRP6, or HBM activity.

2. The method of any of Claims 1, 24, 28, 33, 36, 37, 48, 64, 65, 93, 98, 101, 105, 107, 111, or 112, wherein the Dkk is Dkk-1.

3. The method of any of Claims 1, 24, 28, or 33, wherein the Dkk is Dkk-1 and the Dkk activity is inhibited.

4. The method of Claims 1 or 24, wherein the Dkk activity modulates bone mass and/or lipid levels.

5. The method of Claim 4, wherein bone mass is increased and/or lipid levels are decreased.

6. The method of Claim 5, wherein the increase in bone mass is determined via one or more of a decrease in fracture rate, an increase in bone strength, an increase in bone density, an increase in bone mineral density, an increase in trabecular connectivity, an increase in trabecular density, an increase in cortical density, an increase in bone diameter, and an increase in inorganic bone content.

7. The method of any of Claims 1, 24, 28, or 33, wherein said composition comprises one or more compounds selected from the group consisting of Dkk interacting proteins, or a Dkk-binding fragment thereof.

8. The method of any of Claims 1, 24, 28, or 33, wherein said composition comprises an antisense, a siRNA, or shRNA molecule which recognizes and binds to a nucleic acid encoding one or more Dkk interacting proteins.

9. The method of any of Claims 1, 24, 28, or 33, and wherein said composition comprises a Dkk peptide aptamer.

10. The method of any of Claims 1, 24, 28, or 33, wherein said composition comprises a mimetic of a Dkk peptide aptamer.

11. The method of any of Claims 1, 24, 28, or 33, wherein said composition inhibits Dkk binding to LRP5, LRP6, or HBM.

12. The method of any of Claims 1, 24, 28, or 33, wherein said composition enhances binding of Dkk to LRP5, LRP6, or HBM.

13. The method of any of Claims 1, 24, 28, or 33, wherein said composition comprises a Dkk interacting protein peptide aptamer.

14. The method of any of Claims 1, 24, 28, or 33, wherein said composition comprises a mimetic of a Dkk interacting protein peptide aptamer.

15. The method of any of Claims 1, 24, 28 or 33, wherein said composition inhibits Dkk interacting protein or Dkk-binding fragment thereof binding to Dkk.

16. The method of any of Claims 1, 24, 28, or 33, wherein said composition enhances binding of Dkk interacting protein or Dkk-binding fragment thereof to Dkk.

17. The method of any of Claims 1, 24, 28, or 33, wherein said subject is a vertebrate or an invertebrate organism.

18. The method of any of Claims 1, 24, 28, or 33, wherein said subject is a mammal.

19. The method of any of Claims 1, 24, 28, or 33, wherein said subject is a canine, a feline, an ovine, a primate, an equine, a porcine, a caprine, a camelid, an avian, a bovine, or a rodent.

20. The method of Claim 19, wherein said primate is a human.

21. The method of any of Claims 1, 24, 28, or 33, wherein said composition comprises an LRP5 peptide aptamer.

22. The method of Claim 21, wherein said peptide aptamer is OST262 (SEQ ID NO:208).

23. The method of any of Claims 1, 24, 28 or 33, wherein the composition comprises an LRP5 antibody or an immunologically active fragment thereof.

24. A method of regulating Dkk-Wnt pathway activity in a subject comprising administering a composition which modulates Dkk activity in an amount effective to regulate Dkk-Wnt pathway activity.

25. The method of Claims 24, 101, or 107, wherein the Wnt is one or more of Wnt1-Wnt19.

26. The method of Claim 25, wherein the Wnt is Wnt1, Wnt3, Wnt3a, or Wnt10b.

27. The method of Claim 24 wherein said composition which modulates Dkk activity or modulates Dkk interaction with LRP5/LRP6/HBM is administered in an amount effective to modulate Wnt signaling.

28. A method of modulating bone mass in a subject comprising administering to the subject a composition which modulates Dkk activity or Dkk interaction with LRP5, LRP6, or HBM in an amount effective to modulate bone mass in the subject.

29. The method of Claim 28, wherein bone mass is increased.

30. The method of the previous claim, wherein the increase in bone mass is determined via one or more of a decrease in fracture rate, an increase in bone strength, an increase in bone density, an increase in bone mineral density, an increase in trabecular connectivity, an increase in trabecular density, an increase in cortical density, an increase in bone diameter, and an increase in inorganic bone content.

31. The method of Claims 28 or 36, wherein said subject has a bone mass disorder selected from the group consisting of a bone development disorder, a bone fracture, age-related loss of bone, chrondrodystrophy, drug-induced bone disorder, high bone turnover, hypercalcemia, hyperostosis, osteogenesis imperfecta, osteomalacia, osteomyelitis, osteoporosis, Paget's disease, osteoarthritis, and rickets.

32. The method of Claim 28, wherein the composition which modulates Dkk activity or Dkk interaction with LRP5, LRP6, or HBM is

administered in an amount effective to modulate the amount of trabecular and/or cortical tissue.

33. A method of modulating lipid levels in a subject comprising administering to the subject a composition which modulates Dkk activity or Dkk interaction with LRP5, LRP6, or HBM in an amount effective to modulate lipid levels in the subject.

34. The method of Claim 33, wherein lipid levels are decreased.

35. The method of Claim 33 or 36, wherein the subject has a lipid-modulated disorder and wherein the lipid-modulated disorder is selected from the group consisting of a cardiac condition, atherosclerosis, familial lipoprotein lipase deficiency, familial apoprotein CII deficiency, familial type 3 hyperlipoproteinemia, familial hypercholesterolemia, familial hypertriglyceridemia, multiple lipoprotein-type hyperlipidemia, elevated lipid levels due to dialysis and/or diabetes, and elevated lipid levels of unknown etiology.

36. A method of diagnosing low or high bone mass and/or high or low lipid levels in a subject comprising examining expression of Dkk, LRP5, LRP6, HBM, or and HBM-like variant in the subject and determining whether Dkk, LRP5, LRP6, HBM, or an HBM-like variant is over- or under-expressed to determine whether subject has (a) high or low bone mass and/or (b) has high or low lipid levels.

37. A method of screening for a compound which modulates the interaction of Dkk with LRP5, LRP6, HBM, or a Dkk-binding fragment of LRP5, LRP6, or HBM comprising:

(a) exposing Dkk and a LRP5, LRP6, and/or HBM binding fragment thereof to a compound; and

(b) determining whether said compound modulates Dkk interaction with the LRP5/LRP6/HBM binding fragment.

38. The method of Claim 37, wherein said modulation is determined by whether said compound binds to Dkk or the LRP5, LRP6, or HBM binding fragment thereof.

39. The method of Claim 37, wherein Dkk or a LRP-binding fragment thereof is attached to a substrate.

40. The method of Claim 37, wherein said compound comprises one or more compounds selected from the group consisting of Dkk interacting proteins, or a Dkk-binding fragment thereof.

41. The method of Claim 37 or 48, wherein said compound comprises a Dkk peptide aptamer.

42. The method of Claim 37 or 48, wherein said compound comprises a mimetic of a Dkk peptide aptamer.

43. The method of Claim 37 or 48, wherein said compound comprises a Dkk interacting protein peptide aptamer.

44. The method of Claim 37 or 48, wherein the compound comprises an LRP5 peptide aptamer.

45. The method of Claim 44, wherein the peptide aptamer is OST262 (SEQ ID NO:208).

46. The method of Claim 37 or 48, wherein the compound comprises an LRP5 antibody.

47. The method of Claim 37 or 48, wherein said compound is a mimetic of a Dkk interacting protein peptide aptamer.

48. A method of screening for a compound which modulates the interaction of Dkk with a Dkk interacting protein comprising:

(a) exposing a Dkk interacting protein or a Dkk-binding fragment thereof to a compound; and

(b) determining whether said compound bound to a Dkk interacting protein or the Dkk-binding fragment thereof; and

(c) further determining whether said compound modulates the interaction of Dkk interacting protein and Dkk.

49. The method of Claim 48, wherein the Dkk interacting protein or a Dkk-binding fragment thereof is attached to a substrate.

50. A composition comprising a LRP5, LRP6, or HBM activity-modulating compound and a pharmaceutically acceptable carrier therefor.

51. The composition of Claim 50, wherein said LRP5, LRP6, or HBM activity-modulating compound comprises a compound which binds to Dkk thereby modulating the interaction of Dkk with LRP5, LRP6, or HBM.

52. The composition of Claim 50, wherein said LRP5, LRP6, or HBM modulating compound comprises one or more Dkk interacting proteins and Dkk-binding fragments thereof.

53. The composition of Claim 50, wherein said LRP5, or LRP6, or HBM modulating compound is a monoclonal antibody or an immunologically active fragment thereof which binds to a Dkk interacting protein, or a Dkk-binding fragment thereof.

54. The composition of Claim 53, wherein the monoclonal antibody is human, chimeric, humanized, primatized®, or bispecific.

55. The composition of Claim 50, wherein said LRP5, LRP6, or HBM modulating compound comprises an antisense, a siRNA, or shRNA molecule which recognizes and binds to a nucleic acid encoding one or more Dkk interacting proteins.

56. The composition of Claim 50, wherein said LRP5, LRP6, or HBM modulating compound comprises a Dkk peptide aptamer.

57. The composition of Claim 50, wherein said LRP5, LRP6, or HBM modulating compound comprises a mimetic of a Dkk peptide aptamer.

58. The composition of Claim 50, wherein said LRP5, LRP6, or HBM modulating compound comprises a Dkk interacting protein peptide aptamer.

59. The composition of Claim 50, wherein said LRP5, LRP6, or HBM modulating compound comprises a mimetic of a Dkk interacting protein peptide aptamer.

60. The composition of Claim 50, wherein the compound comprises an LRP5 peptide aptamer.

61. The composition of Claim 60, wherein the peptide aptamer is OST262.

62. The composition of Claim 50, wherein the compound comprises an LRP5 antibody.

63. A pharmaceutical composition comprising a compound which modulates Dkk activity and a pharmaceutically acceptable carrier therefor.

64. A method for identifying compounds which modulate Dkk and LRP5/LRP6/HBM interactions comprising:

(a) creating an LRP5, LRP6, or HBM fluorescent fusion protein using a first fluorescent tag; and

(b) creating a Dkk fusion protein comprising a second fluorescent tag;

(c) adding a test compound; and

(d) assessing changes in the ratio of fluorescent tag emissions using Fluorescence Resonance Energy Transfer (FRET) or Bioluminescence Resonance Energy Transfer (BRET) to determine whether the compound modulates Dkk and LRP5/LRP6/HBM interactions.

65. A method of identifying binding partners for a Dkk protein comprising the steps of:

(a) exposing the Dkk protein(s) or a LRP5/LRP6 binding fragment thereof to a potential binding partner; and

(b) determining if the potential binding partner binds to a Dkk protein or the LRP5/LRP6 binding fragment thereof.

66. A nucleic acid encoding a Dkk interacting protein peptide aptamer comprising a nucleic acid encoding a scaffold protein in-frame with the activation

domain of Gal4 or LexA that is in-frame with a nucleic acid that encodes a Dkk interacting protein amino acid sequence.

67. A vector comprising the nucleic acid of Claim 66.

68. The nucleic acid of Claim 66, wherein the scaffold protein is trxA.

69. A method of detecting a modulatory activity of a compound on the binding interaction of a first peptide and a second peptide of a peptide binding pair that bind through extracellular interaction in their natural environment, comprising:

(i) culturing at least one eukaryotic cell comprising:

a) a nucleotide sequence encoding a first heterologous fusion protein comprising the first peptide or a segment thereof joined to a transcriptional activation protein DNA binding domain;

b) a nucleotide sequence encoding a second heterologous fusion protein comprising the second peptide or a segment thereof joined to a transcriptional activation protein transcriptional activation domain;

wherein binding of the first peptide or segment thereof and the second peptide or segment thereof reconstitutes a transcriptional activation protein; and

c) a reporter element activated under positive transcriptional control of the reconstituted transcriptional activation protein, wherein expression of the reporter element produces a selected phenotype;

(ii) incubating the eukaryotic cell in the presence of a compound under conditions suitable to detect the selected phenotype; and

(iii) detecting the ability of the compound to affect the binding interaction of the peptide binding pair by determining whether the compound affects the expression of the reporter element which produces the selected phenotype;

wherein (1) said first peptide is a Dkk peptide and the second peptide is a peptide selected from LRP5, HBM, LRP6 and the Dkk-binding portion of LRP5/LRP6/HBM or (2) said first peptide is a Dkk interacting protein or the Dkk-binding fragment thereof and said second peptide is a Dkk peptide.

70. The method of Claim 69, wherein the eukaryotic cell is a yeast cell.

71. The method of Claim 70, wherein the yeast cell is *Saccharomyces*.

72. The method of Claim 71, wherein the *Saccharomyces* cell is *Saccharomyces cerevisiae*.

73. The method of Claim 69, wherein the Dkk is Dkk-1 and wherein the compound comprises one or more Dkk interacting proteins or a Dkk-binding fragment thereof.

74. The method of Claim 73, wherein the compound is directly added to assay.

75. The method of Claim 73, wherein the compound is recombinantly expressed by said eukaryotic cell in addition to said first and second peptides.

76. The method of Claim 69, wherein the compound comprises a Dkk peptide aptamer.

77. The method of Claim 69, wherein the compound comprises a mimetic of a Dkk peptide aptamer.

78. The method of Claim 69, wherein the compound comprises a Dkk interacting protein peptide aptamer.

79. The method of Claim 69, wherein the compound comprises a mimetic of a Dkk interacting protein peptide aptamer.

80. The method of Claim 69, wherein the eukaryotic cell further comprises at least one endogenous nucleotide sequence selected from the group consisting of a nucleotide sequence encoding the DNA binding domain of a transcriptional activation protein, a nucleotide sequence encoding the transcriptional activation domain of a transcriptional activation protein, and a nucleotide sequence encoding the reporter element, wherein at least one of the endogenous nucleotide sequences is inactivated by mutation or deletion.

81. The method of Claim 69, wherein the peptide binding pair comprises a ligand and a receptor to which the ligand binds.

82. The method of Claim 69, wherein the transcriptional activation protein is Gal4, Gcn4, Hap1, Adr1, Swi5, Ste12, Mcm1, Yap1, Ace1, Ppr1, Arg81, Lac9, Qa1F, VP16, or a mammalian nuclear receptor.

83. The method of Claim 69, wherein at least one of the heterologous fusion proteins is expressed from an autonomously-replicating plasmid.

84. The method of Claim 69, wherein the DNA binding domain is a heterologous DNA-binding domain of a transcriptional activation protein.

85. The method of Claim 84, wherein the DNA binding protein is selected from the group consisting of a mammalian steroid receptor and bacterial LexA protein.

86. The method of Claim 69, wherein the reporter element is selected from the group consisting of *lacZ*, a polynucleotide encoding luciferase, a polynucleotide encoding green fluorescent protein (GFP), and a polynucleotide encoding chloramphenicol acetyltransferase.

87. The method of Claim 86, wherein the reporter element is LacZ.

88. The method of Claim 69, wherein the test sample comprises an LRP5 peptide aptamer.

89. The method of Claim 88, wherein the peptide aptamer is OST262 (SEQ ID NO:208).

90. The method of Claim 69, wherein the test sample comprises an LRP5 antibody.

91. A transgenic animal wherein Dkk-1 is knocked out in a tissue-specific fashion.

92. The transgenic animal of Claim 91, wherein the tissue specificity is bone tissue, cancer tissue, or liver tissue.

93. A method for identifying potential compounds which modulate Dkk activity comprising:

a) measuring the effect on binding of one or more Dkk interacting proteins, or a Dkk-binding fragment thereof, with Dkk or a

fragment thereof in the presence and absence of a compound; and

b) identifying as a potential Dkk modulatory compound a compound which modulates the binding between one or more Dkk interacting proteins or Dkk-binding fragment thereof and Dkk or fragment thereof.

94. A peptide aptamer of Figure 3 (SEQ ID NOs:171-188) or Figure 4 (SEQ ID NOs:189-192).

95. An antibody or antibody fragment which recognizes and binds to one or more peptides of amino acid sequences GNKYQTIDNYQPYPC (SEQ ID NO:118), LDGYSRRTTLSSKMYHTKGQEG (SEQ ID NO:119), RIQKDHHQASNSSRLHTCQRH (SEQ ID NO:120), RGEIEETITESFGND (SEQ ID NO:121), EIFQRCYCGEGLSCRIQKD (SEQ ID NO:122), MYWTDWVETPRIE (SEQ ID NO:123), MYWTDWGETPRIE (SEQ ID NO:124), KRTGGKRKEILSA (SEQ ID NO:125), ERVEKTTGDKRTRIQGR (SEQ ID NO:126), KQQCDSFPDCIDGSDE (SEQ ID NO:127), or a Dkk-1 amino acid sequence selected from the group consisting Asn34-His266 (SEQ ID NO:110), Asn34-Cys245 (SEQ ID NO:111), Asn34-Lys182 (SEQ ID NO:112), Cys97-His266 (SEQ ID NO:113), Val139-His266 (SEQ ID NO:114), Gly183-His266 (SEQ ID NO:115), Cys97-Cys245 (SEQ ID NO:116), or Val139-Cys245 (SEQ ID NO:117).

96. The antibody or antibody fragment of Claim 95, wherein the antibody is a monoclonal antibody.

97. The antibody or antibody fragment of Claim 95, wherein the antibody is a polyclonal antibody

98. A method of identifying Dkk interacting proteins which modulate the interaction of Dkk with the Wnt signaling pathway comprising:

(a) injecting Dkk and potential Dkk interacting protein mRNA into a *Xenopus* blastomere; and

(b) assessing axis duplication or analyzing marker gene expression; and

(c) identifying compositions which elicit changes in axis duplication or marker gene expression as Dkk interacting proteins which modulate the interaction of Dkk with the Wnt signaling pathway.

99. The method of Claim 98, wherein the mRNA of HBM, LRP5/6, any Wnt, Wnt antagonist, Wnt pathway modulator, or combination of these is co-injected into the *Xenopus* blastomere.

100. The method of Claim 98, wherein the marker gene analyzed is Siamois, Xnr3, slug, Xbra, HNK-1, endodermin, Xlhbox8, BMP2, BMP4, XLRP6, EF-1, or ODC.

101. A method for identifying Dkk interacting proteins which modulate the interaction of Dkk with the Wnt signaling pathway comprising:

(a) transfecting cells with constructs containing Dkk and potential Dkk interacting proteins; and

(b) assessing changes in expression of a reporter gene linked to a Wnt-responsive promoter; and

(c) identifying as a Dkk interacting protein any protein which alters reporter gene expression compared with cells transfected with a Dkk construct alone.

102. The method of Claim 101, wherein the cells are HOB-03-CE6, HEK293, or U2OS cells.

103. The method of Claim 101, wherein the Wnt-responsive promoter is TCF or LEF.

104. The method of Claim 101, wherein the cells are co-transfected with CMV β-galactosidase.

105. A method for identifying compounds which modulate Dkk and LRP5/LRP6/HBM interactions comprising:

(a) immobilizing LRP5/LRP6/HBM to a solid surface; and

(b) treating the solid surface with a secreted Dkk protein or a secreted epitope-tagged Dkk and a test compound; and

(c) determining whether the compound regulates binding between Dkk and LRP5/LRP6/HBM using antibodies to Dkk or the epitope tag or by directly measuring activity of an epitope tag.

106. The method of Claim 105, wherein the epitope tag is alkaline phosphatase, histidine, or a V5 tag.

107. A method for identifying compounds which modulate the interaction of Dkk with the Wnt signaling pathway comprising:

(a) transfecting cells with constructs containing Dkk and Wnt proteins;

(b) assessing changes in expression of a reporter element linked to a Wnt- responsive promoter; and

(c) identifying as a Dkk/Wnt interaction modulating compound any compound which alters reporter gene expression compared with cells transfected with a Dkk construct alone.

108. The method according to Claim 107, wherein Wnt3a and Wnt1 constructs are co-transfected into the cells.

109. The method according to Claim 107, wherein the cells are U2-OS, HOB-03-CE6, or HEK293 cells.

110. The method according to Claim 107, wherein the reporter element used is TCF-luciferase, tk-Renilla, or a combination thereof.

111. A method of testing compounds that modulate Dkk-mediated activity in a mammal comprising

(a) providing a group of transgenic animals having (1) a regulatable one or more Dkk genes, (2) a knock-out of Dkk genes, or (3) a knock-in of one or more Dkk genes;

(b) providing a second group of control animals respectively for the group of transgenic animals in step (a); and

(c) exposing the transgenic animal group and control animal group to a potential Dkk-modulating compound which modulates bone mass or lipid levels; and

(d) comparing the transgenic animals and the control group of animals and determining the effect of the compound on bone mass or lipid levels in the transgenic animals as compared to the control animals.

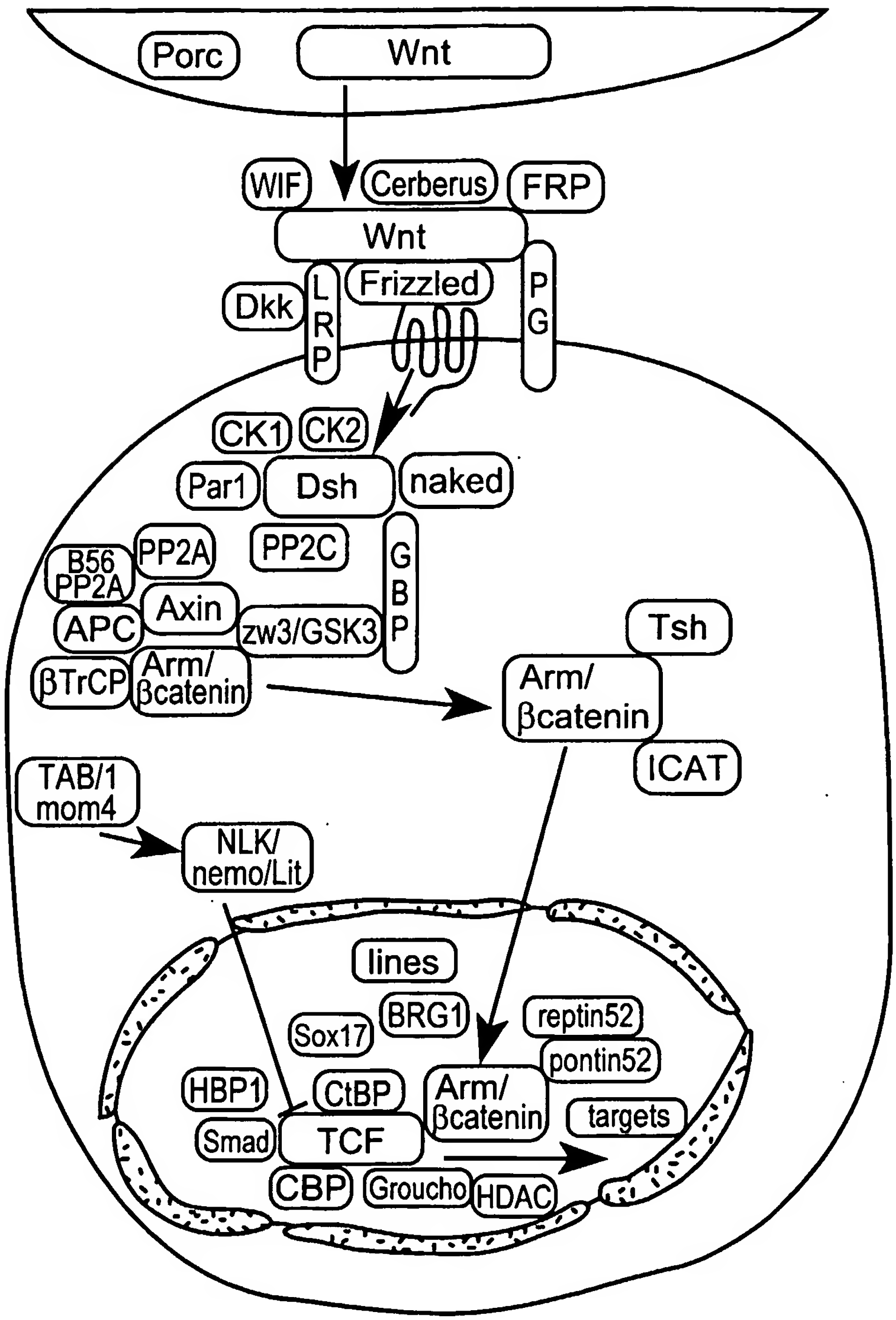
112. A method of screening for compounds or compositions which modulate the interaction of Dkk and a Dkk interacting protein comprising:

(a) exposing a Dkk interacting proteins or a Dkk-binding fragment thereof to a compound; and

(b) determining whether said compound binds to a Dkk interacting proteins or the Dkk-binding fragment thereof.

113. The method of Claim 112, wherein said modulation is determined by whether said compound binds to the Dkk interacting protein or the Dkk-binding fragment thereof.

114. An antibody or antibody fragment which recognizes and binds to a sequence depicted in Figure 3 (SEQ ID NOs:171-188) or Figure 4 (SEQ ID NOs:189-192).



Model of Wnt signaling

FIG. 1

Sequence of baits used in Y2H screens
>DKK1 (SEQ ID N0: 168)

AATTCCAACGCTATCAAGAACCTGCCCCCACCGCTGGGCGGCGCTG
CGGGGCACCCAGGCTCTGCAGTCAGCGCCGCGCCGGGAATCCTGTA
CCCGGGCGGGAATAAGTACCAGACCATTGACAACTACCAGCCGTAC
CCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCT
AGTCCCACCCGCGGAGGGGACGCGGGCGTGCAAATCTGTCTCGCCT
GCAGGAAGCGCCGAAAACGCTGCATGCGTCACGCTATGTGCTGCCC
CGGGAATTACTGCAAAAATGGAATATGTGTGTCTTCTGATCAAAAT
CATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTA
ATGATCATAGCACCTTGGATGGGTATTCCAGAAGAACCACCTTGTC
TTCAAAAATGTATCACACCAAAGGACAAGAAGGTTCTGTTTGTCTC
CGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTAGACACTTCTG
GTCCAAGATCTGTAAACCTGTCCTGAAAGAAGGTCAAGTGTGTACC
AAGCATAGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTT
GTTACTGTGGAGAAGGTCTGTCTTGCCGGATACAGAAAGATCACCA
TCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTCAGAGACACTAA

FIG. 2A

>zmax1 LBD1 (SEQ ID NO: 169)

CTCATCCTGCCCCTGCATGGACTGAGGAACGTCAAAGCCATCGACTAT
GACCCACTGGACAAGTTCATCTACTGGGTGGATGGGCGCCAGAACATC
AAGCGAGCCAAGGACGACGGGACCCAGCCCTTTGTTTTGACCTCTCTG
AGCCAAGGCCAAAACCCAGACAGGCAGCCCCACGACCTCAGCATCGA
CATCTACAGCCGGACACTGTTCTGGACGTGCGAGGCCACCAATACCAT
CAACGTCCACAGGCTGAGCGGGGAAGCCATGGGGGTGGTGCTGCGTG
GGGACCGCGACAAGCCCAGGGCCATCGTCGTCAACGCGGAGCGAGGG
TACCTGTACTTCACCAACATGCAGGACCGGGCAGCCAAGATCGAACGC
GCAGCCCTGGACGGCACCGAGCGCGAGGTCCTCTTCACCACCGGCCTC
ATCCGCCCTGTGGCCCTGGTGGTAGACAACACACTGGGCAAGCTGTTC
TGGGTGGACGCGGACCTGAAGCGCATTGAGAGCTGTGACCTGTCAGG
GGCCAACCGCCTGACCCTGGAGGACGCCAACATCGTGCAGCCTCTGGG
CCTGACCATCCTTGGCAAGCATCTCTACTGGATCGACCGCCAGCAGCA
GATGATCGAGCGTGTGGAGAAGACCACCGGGGACAAGCGGACTCGCA
TCCAGGGCCGTGTCGCCCACCTCACTGGCATCCATGCAGTGGAGGAAG
TCAGCCTGGAGGAGTTCTCAGCCCACCCATGTGCCCGTGACAATGGTG
GCTGCTCCCACATCTGTATTGCCAAGGGTGATGGGACACCACGGTGCT
CATGCCCAGTCCACCTCGTGCTCCTGCAGAACCTGCTGACCTGTGGAG
AGCCGCCCACCTGCTCCCCGGACCAGTTTGCATGTGCCACAGGGGAGA
TCGACTGTATCCCCGGGGCCTGGCGCTGTGACGGCTTTCCCGAGTGCG
ATGACCAGAGCGACGAGGAGGGCTGCCCCGTGTGCTCCGCCGCCCAGT
TCCCCTGCGCGCGGGGTCAGTGTGTGGACCTGCGCCTGCGCTGCGACG
GCGAGGCAGACTGTCAGGACCGCTCAGACGAGGCGGACTGTGACGCC
ATCTGCCTGCCCAACCAGTTCCGGTGTGCGAGCGGCCAGTGTGTCCTC
ATCAAACAGCAGTGCGACTCCTTCCCCGACTGTATCGACGGCTCCGAC
GAGCTCATGTGTGAAATCACCAAGCCGCCC

FIG. 2B

>zmax1 LBD4 (SEQ ID NO: 170)

AGGGCCATCGTCGTCAACGCGGAGCGAGGGTACCTGTACTTCACCAA
CATGCAGGACCGGGCAGCCAAGATCGAACGCGCAGCCCTGGACGGCA
CCGAGCGCGAGGTCCTCTTCACCACCGGCCTCATCCGCCCTGTGGCCC
TGGTGGTAGACAACACACTGGGCAAGCTGTTCTGGGTGGACGCGGAC
CTGAAGCGCATTGAGAGCTGTGACCTGTCAGGGGCCAACCGCCTGAC
CCTGGAGGACGCCAACATCGTGCAGCCTCTGGGCCTGACCATCCTTGG
CAAGCATCTCTACTGGATCGACCGCCAGCAGCAGATGATCGAGCGTG
TGGAGAAGACCACCGGGGACAAGCGGACTCGCATCCAGGGCCGTGTC
GCCCACCTCACTGGCATCCATGCAGTGGAGGAAGTCAGCCTGGAGGA
GTTCTCAGCCCACCCATGTGCCCGTGACAATGGTGGCTGCTCCCACAT
CTGTATTGCCAAGGGTGATGGGACACCACGGTGCTCATGCCCAGTCCA
CCTCGTGCTCCTGCAGAACCTGCTGACCTGTGGAGAGCCGCCCACCTG
CTCCCCGGACCAGTTTGCATGTGCCACAGGGGAGATCGACTGTATCCC
CGGGGCCTGGCGCTGTGACGGCTTTCCCGAGTGCGATGACCAGAGCG
ACGAGGAGGGCTGCCCCGTGTGCTCCGCCGCCCAGTTCCCCTGCGCGC
GGGGTCAGTGTGTGGACCTGCGCCTGCGCTGCGACGGCGAGGCAGAC
TGTCAGGACCGCTCAGACGAGGCGGACTGTGACGCCATCTGCCTGCC
CAACCAGTTCCGGTGTGCGAGCGGCCAGTGTGTCCTCATCAAACAGC
AGTGCGACTCCTTCCCCGACTGTATCGACGGCTCCGACGAGCTCATGT
GTGAAATCACCAAGCCGCCCTAAGCGGCCGC

FIG. 2C

Screen of DKK1 X Peptide Library

name	motif	# hits	SEQ ID NO:
252-1	SVGCLLCAGLGVWSLS	3	171
252-2	WCCCGLFRGVCVWSCGAD D	2	172
252-3	GWRRCDWCGCVSWCWV	1	173
252-4	MPGSVSHCWGGICEAL	8	174
252-15	SCCAVDVCLRCGGWFR	1	175
252-16	SVLGTCCCCGGWILCE	2	176
252-17	VLSVCEVCGGVFVRRC	1	177
252-18	GMWYWSGRDCALCWL	1	178
252-19	CTAVMWGVGSVAYLGE	1	179
252-20	WCWWCGCRGVVWR	1	180
252-21	CVCASFCCCVCGLRLL	1	181
252-23	TYEVCEECGGRVRMWV	6	182
252-25	VVVCASCGQVWHGSGA	2	183
252-26	CCRCCHCWDCEWHMCV	1	184
252-27	FCASCCWCGCDCFGWV	2	185
252-32	CDYCWSCGVWCPSSWL	3	186
252-47	VYLCVWCGAARFGCYG	1	187
252-48	FCVCGCCWCWCAACWC	1	188

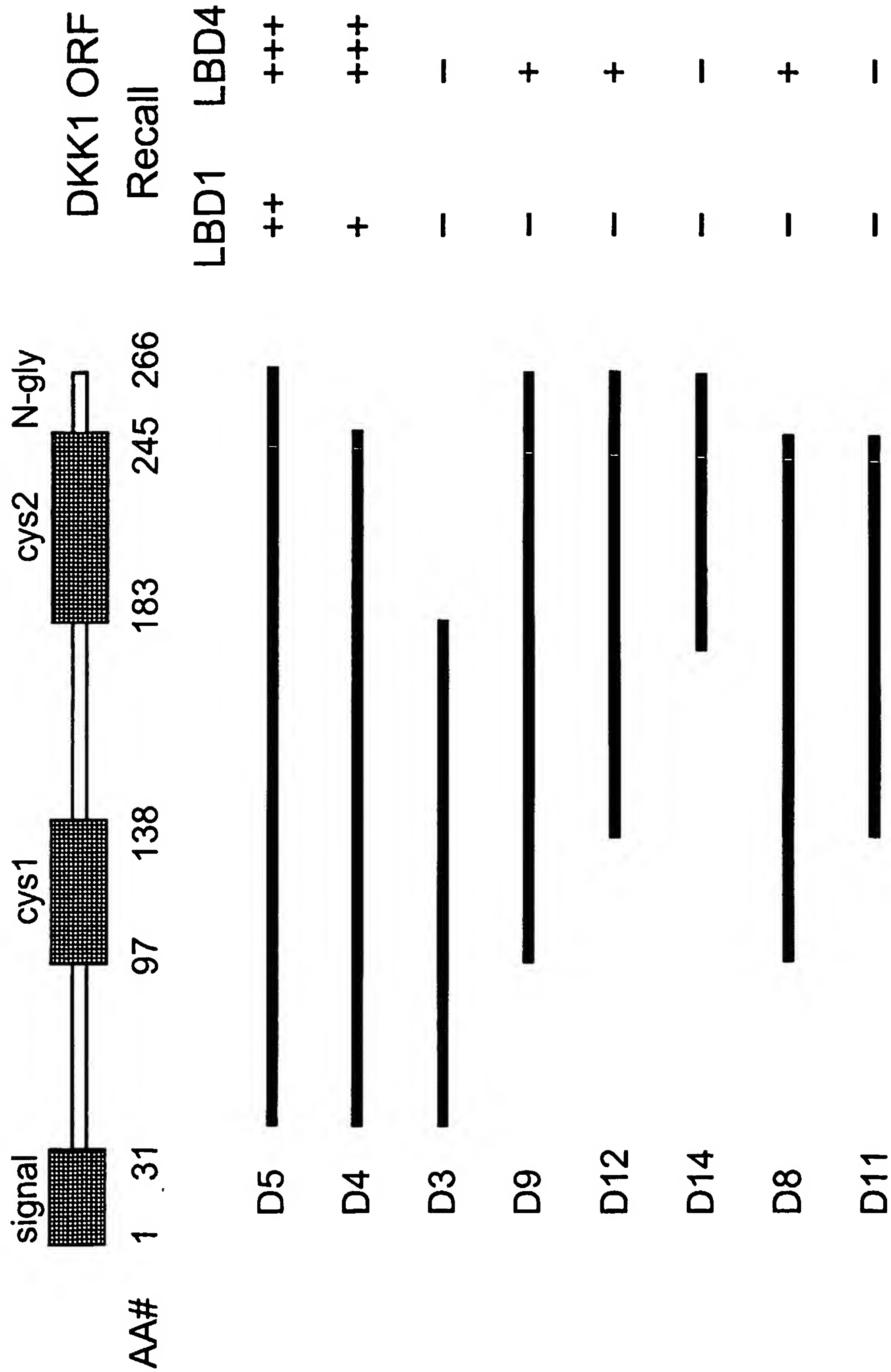
FIG. 3

peptide #	peptide seq	# hits	SEQ ID NO:
9	VVLCSRCGRLWRWSCG	1	189
12	EVRQVTCIRCRRGFLL	1	190
13	GGGGMWEAWSCYACG	1	191
14	GWRWCGRCGALWWRRV	3	192

FIG. 4

Gene	Genbank Accession #	Protein Accession #
granulin	M75161	AAA58617
similar to cys/His rich protein	BC004544	AAH04544
IGF-BINDING PROTEIN 6	M69054	AAA88070
latent TGFb binding protein 4	AF051344	AAC39879
NOTCH 2	AF315356	AAG37073
fibulin 1	X53743	CAA37772
MDC15 (ADAM15)	U46005	AAC51112
DKFZp761G02121(notch1 Ca++ binding like)	AL137311	CAB70690
chordin	AF076612	AAC69835
fibronectin 1	U42594	AAD00019
MG50(melanoma associated antigen)	AF200348	AAF06354
unknown (notch 4-like)	AX068260	CAC27245
Slit 1	AB017167	BAA35184
tomoregulin (agarin repeat homology)	AB004064	BAA90820
sprouty 1	AF041037	AAC39566
sprouty 2	AF039843	AAC04258
NOV1	X96584	CAA65403
agrin	AF016903	AAC39776
fibrillin 1	L13923	AAB02036
thrombospondin1	X04665	CAA28370
ADAM19	AF134707	AAF22162
Nafl alpha	AJ011895	CAA09855
laminin alpha 5	Z95636	CAB09137
CRIM1	AF167706	AAF34409
nidogen	M30269	AAA59932
fibulin-2	X82494	CAA57876
thrombospondin 2	L12350	AAA03703
KIAA1323	AB037744	BAA92561
fibrillin-2	U03272	AAA18950
MEGF9	AB011542	BAA32470
integrin beta 1	X07979	CAA30790
matrilin-2 precursor	U69263	AAC51260
tenascin	X56160	A32160

FIG. 5



DKK1 ORF
Recall
LBD1 LBD4
AA#
signal
cys1
cys2
N-gly
1
31
97
138
183
245
266
D5 ++ +++
D4 + +++
D3 - -
D9 - +
D12 - +
D14 - -
D8 - +
D11 - -

FIG. 6

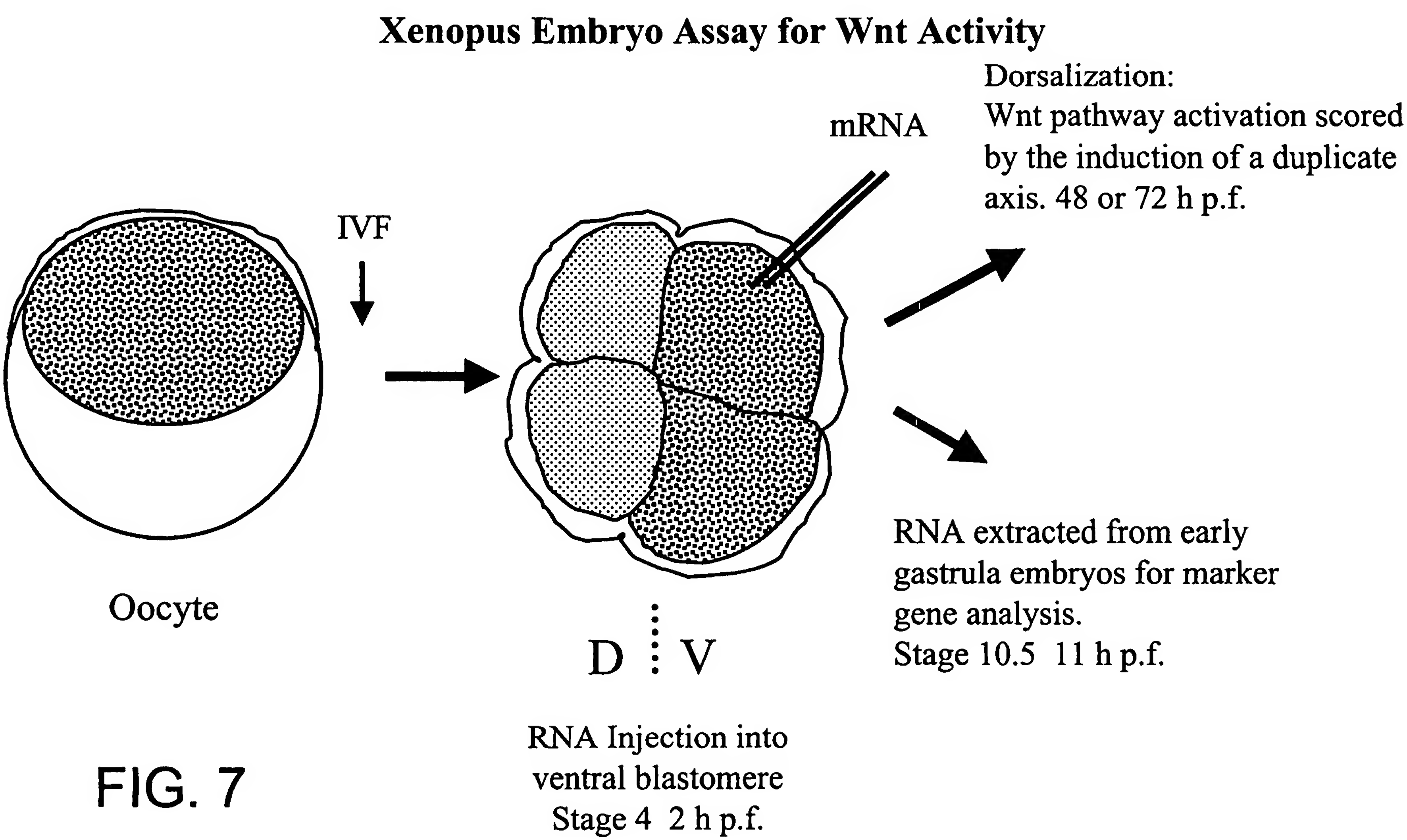

Xenopus Embryo Assay for Wnt Activity
IVF
mRNA
Dorsalization:
Wnt pathway activation scored by the induction of a duplicate axis. 48 or 72 h p.f.
RNA extracted from early gastrula embryos for marker gene analysis.
Stage 10.5 11 h p.f.
Oocyte
D
V
RNA Injection into ventral blastomere
Stage 4 2 h p.f.

FIG. 7

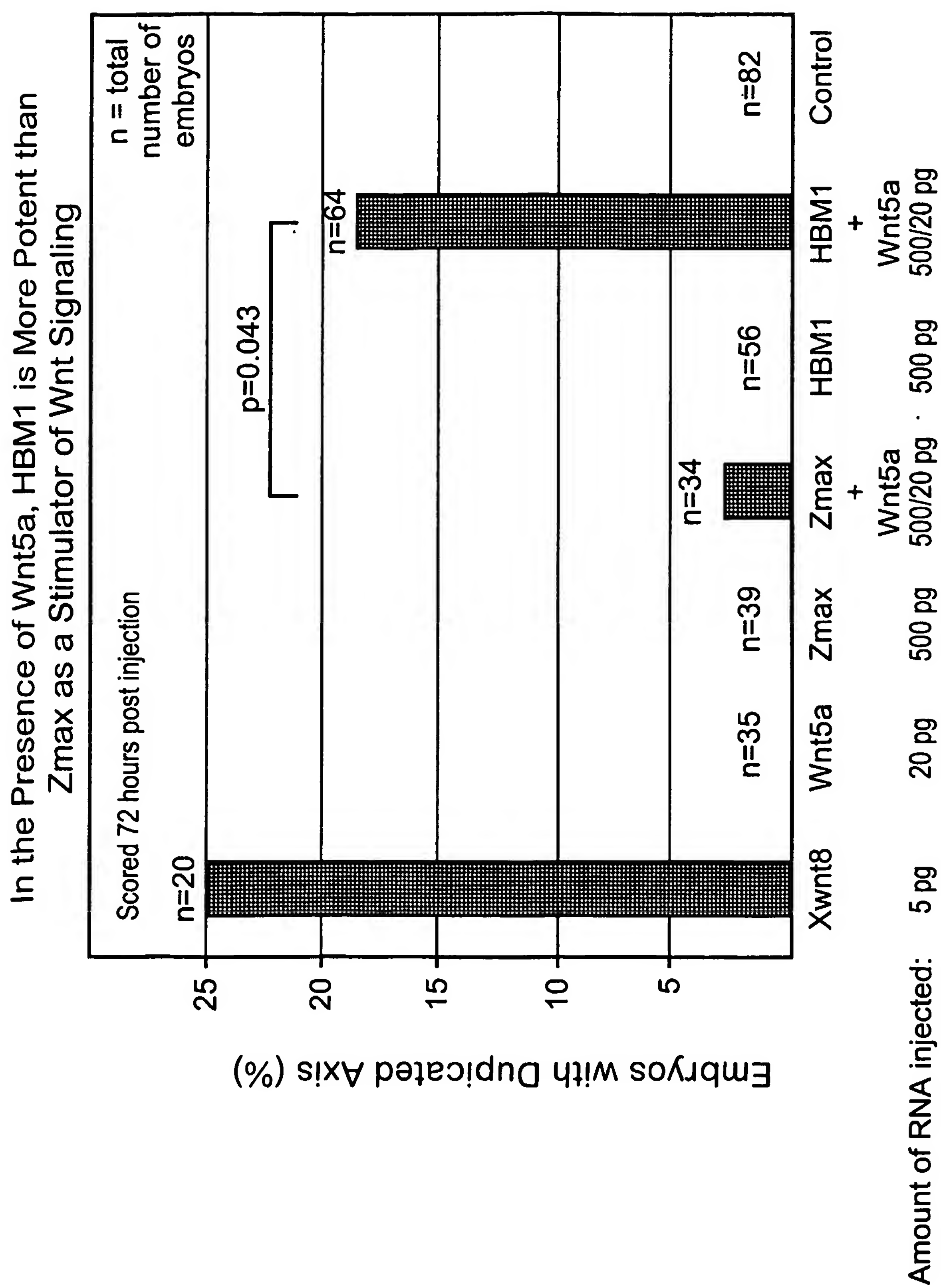

In the Presence of Wnt5a, HBM1 is More Potent than Zmax as a Stimulator of Wnt Signaling
Scored 72 hours post injection
n = total number of embryos
Embryos with Duplicated Axis (%)
25
20
15
10
5
n=20
n=35
n=39
n=34
n=56
n=64
n=82
p=0.043
Xwnt8
Wnt5a
Zmax
Zmax + Wnt5a
HBM1
HBM1 + Wnt5a
Control
Amount of RNA injected:
5 pg
20 pg
500 pg
500/20 pg
500 pg
500/20 pg

FIG. 8

Both Zmax and HBM1, in the presence of Wnt5a, induce secondary axis formation in Xenopus (photos at 48 hrs post-injection)

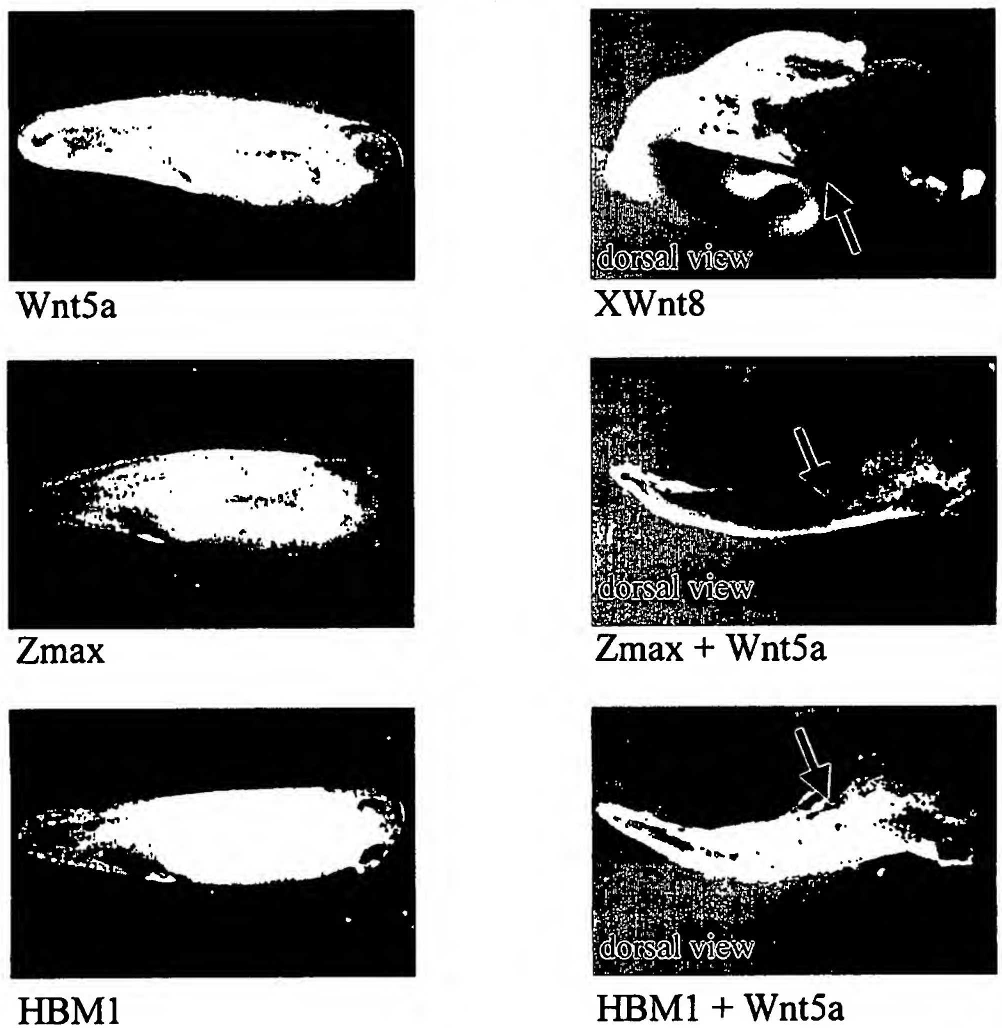
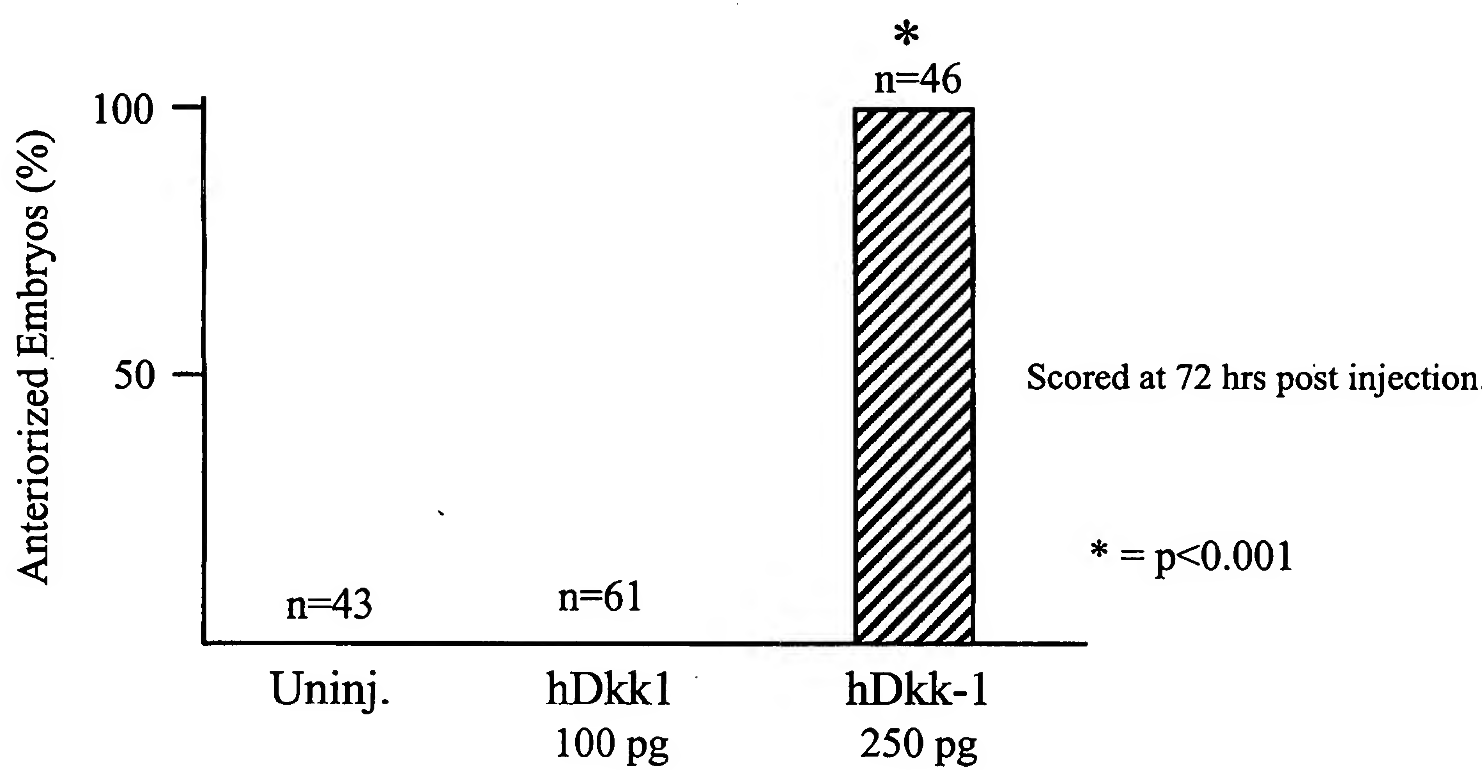


FIG. 9

Human Dkk-1 Represses **the Canonical Wnt Pathway**



•Confirms our human construct is active.
•Reproduces reported dose-response.

FIG. 10

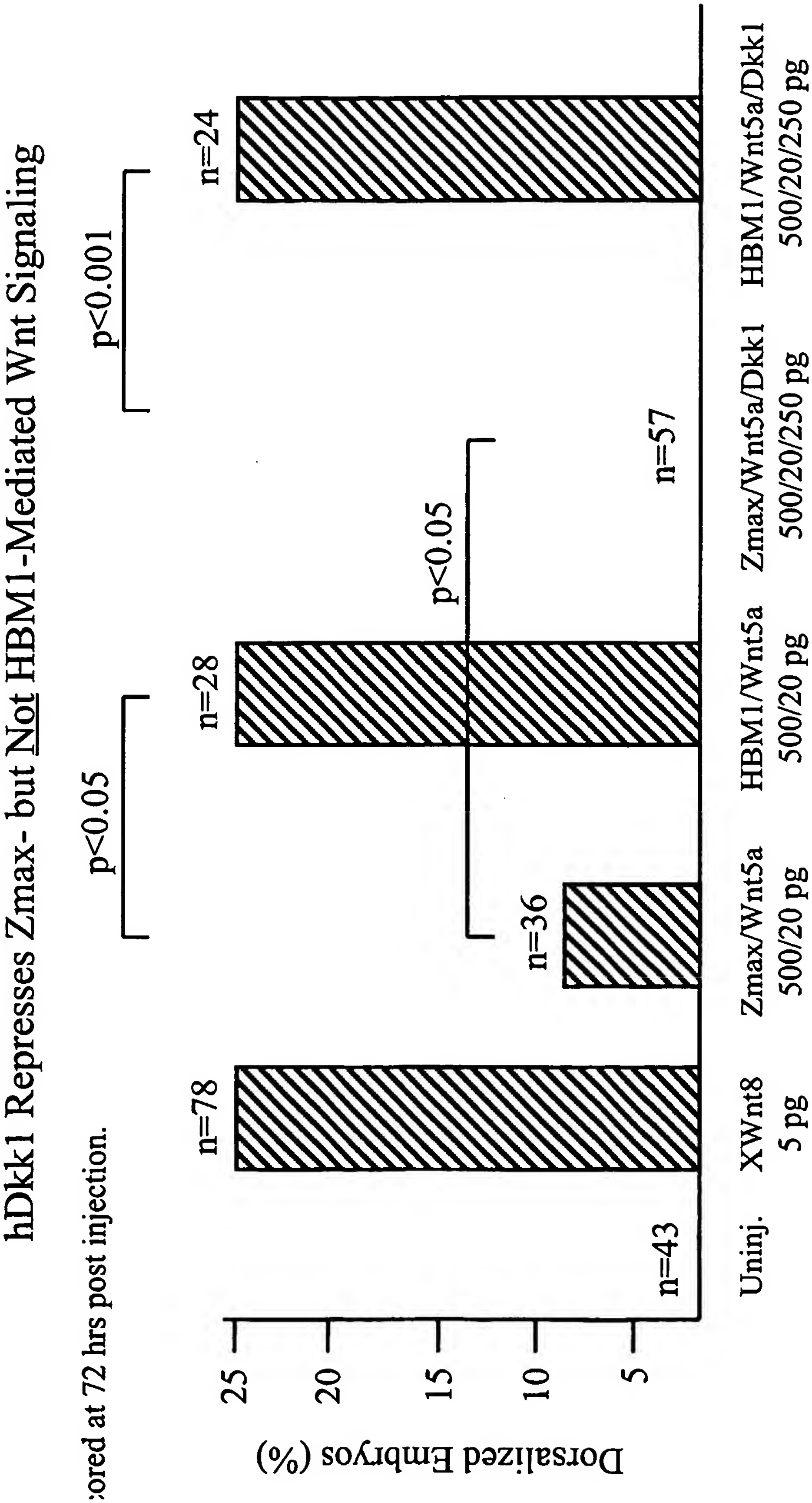

hDkk1 Represses Zmax- but Not HBM1-Mediated Wnt Signaling
:ored at 72 hrs post injection.
Dorsalized Embryos (%)
25
20
15
10
5
n=43
n=78
n=36
n=28
n=57
n=24
p<0.05
p<0.05
p<0.001
Uninj.
XWnt8
5 pg
Zmax/Wnt5a
500/20 pg
HBM1/Wnt5a
500/20 pg
Zmax/Wnt5a/Dkk1
500/20/250 pg
HBM1/Wnt5a/Dkk1
500/20/250 pg

FIG. 11

Listed are the pcDNA3.1 construct names followed by the DNA sequence
OST258 (control for OST 259-OST262 and OST264,OST265)
AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTTCCAGGTTCCACTGGTGACGG
ATCC

OST259 (SEQ ID NO: 193)
AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTTCCAGGTTCCACTGGTGACGG
ATCCATGAGCGATAAAATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGGGGCGATCC
TCGTCGATTTCTGGGCAGAGTGGTGCGGTCCGAATTCCGTGGTTCTGTGTTCGCGTTGTGGGCGTTTGTGGCGGTGG
TCGTGTGGGACTAGTGGTCCGTGCAAAATGATCGCCCCGATTCTGGATGAAATCGCTGACGAATATCAGGGCAAACT
GACCGTTGCAAAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGC
TGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGAC
GCTAACCTGGCGTAAGCGGCCGC

OST260 (SEQ ID NO: 194)
AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTTCCAGGTTCCACTGGTGACGG
ATCCATGAGCGATAAAATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGGGGCGATCC
TCGTCGATTTCTGGGCAGAGTGGTGCGGTCCGAATTCCGGGTGGCGGTGGTGTGGTCGGTGTGGGGCTTTGTGGTGG
CGGCGTGTTACTAGTGGTCCGTGCAAAATGATCGCCCCGATTCTGGATGAAATCGCTGACGAATATCAGGGCAAACT
GACCGTTGCAAAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGC
TGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGAC
GCTAACCTGGCGTAAGCGGCCGC

OST261 (SEQ ID NO: 195)
AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTTCCAGGTTCCACTGGTGACGG
ATCCATGAGCGATAAAATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGGGGCGATCC
TCGTCGATTTCTGGGCAGAGTGGTGCGGTCCGAATTCCGAGGTGCGGCAGGTTACGTGTATTAGGTGTCGTCGGGGT
TTTCTGTTGACTAGTGGTCCGTGCAAAATGATCGCCCCGATTCTGGATGAAATCGCTGACGAATATCAGGGCAAACT
GACCGTTGCAAAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGC
TGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGAC
GCTAACCTGGCGTAAGCGGCCGC

OST262 (SEQ ID NO: 196)
AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTTCCAGGTTCCACTGGTGACGG
ATCCATGAGCGATAAAATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGGGGCGATCC
TCGTCGATTTCTGGGCAGAGTGGTGCGGTCCGAATTCCGGTGGTGGGGGGATGATTTGGGAGGCTTGGAGTTGTTAT
GCGTGTGGGACTAGTGGTCCGTGCAAAATGATCGCCCCGATTCTGGATGAAATCGCTGACGAATATCAGGGCAAACT
GACCGTTGCAAAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGC
TGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGAC
GCTAACCTGGCGTAAGCGGCCGC

OST263 (SEQ ID NO: 197)
AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTTCCAGGTTCCACTGGTGACGG
ATCCATGAGCGATAAAATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGGGGCGATCC
TCGTCGATTTCTGGGCAGAGTGGTGCGGTCCGAATTCCTTGTGGATTGGGCCGGGTGATCAGGGTCTGTTTCGGCGT
TTTGTTTTTACTAGTGGTCCGTGCAAAATGATCGCCCCGATTCTGGATGAAATCGCTGACGAATATCAGGGCAAACT
GACCGTTGCAAAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGGTATCCCGACTCTGC
TGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGAAAGAGTTCCTCGAC
GCTAACCTGGCGTAAGCGGCCGC

FIGURE 12A

OST264 (SEQ ID NO: 198)
AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTTCCAGGTTCCACTGGTGACGG
ATCCGTGTCTTCTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGATCATA
GCACCTTGGATGGGTATTCCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAAAGGACAAGAAGGTTCTGTT
TGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTAGACACTTCTGGTCCAAGATCTGTAAACCTGTCCT
GAAAGAAGGTCAAGTGTGTACCAAGCATAGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTGTTACTGTG
GAGAAGGTCTGTCTTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTCAGAGA
CACTAAGCGGCCGC

OST265 (SEQ ID NO: 199)
AAGCTTGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTTCCAGGTTCCACTGGTGACGG
ATCCTGCGCTAGTCCCACCCGCGGAGGGGACGCGGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACGCT
GCATGCGTCACGCTATGTGCTGCCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTTCTGATCAAAATCATTTC
CGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGATCATAGCACCTTGGATGGGTATTCCAGAAGAAC
CACCTTGTCTTCAAAAATGTATCACACCAAAGGACAAGAAGGTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAG
GATTGTGTTGTGCTAGACACTTCTGGTCCAAGATCTGTAAACCTGTCCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTGTTACTGTGGAGAAGGTCTGTCTTGCTAAGCGGCCGC

OST266 (SEQ ID NO: 200)
AAGCTTGCCACCATGGGCGATAAAATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGG
GGCGATCCTCGTCGATTTCTGGGCAGAGTGGTGCGGTCCGAATTCCTATGCGTGGTTGTTTTCTTGTAGTAGGTGTA
GGTGGTGGTTGCCTTGGACTAGTGGTCCGTGCAAAATGATCGCCCCGATTCTGGATGAAATCGCTGACGAATATCAG
GGCAAACTGACCGTTGCAAAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGGTATCCC
GACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGAAAGAGT
TCCTCGACGCTAACCTGGCGTAAGCGGCCGC

OST267 (SEQ ID NO: 201)
AAGCTTGCCACCATGGGCGATAAAATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGG
GGCGATCCTCGTCGATTTCTGGGCAGAGTGGTGCGGTCCGAATTCCATTTGTGAGGTTGTGAGGTTGTGGAGTCGGT
ATCCTTGGTCTTGGGTGACTAGTGGTCCGTGCAAAATGATCGCCCCGATTCTGGATGAAATCGCTGACGAATATCAG
GGCAAACTGACCGTTGCAAAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGGTATCCC
GACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGAAAGAGT
TCCTCGACGCTAACCTGGCGTAAGCGGCCGC

OST268 (SEQ ID NO: 202)
AAGCTTGCCACCATGGGCGATAAAATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGG
GGCGATCCTCGTCGATTTCTGGGCAGAGTGGTGCGGTCCGAATTCCGGTTGTACTAGTGCGGTGTGTGGTGCTTGGG
CTGAGGCGGGTAGGTTTTATTGTACTAGTGGTCCGTGCAAAATGATCGCCCCGATTCTGGATGAAATCGCTGACGAA
TATCAGGGCAAACTGACCGTTGCAAAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGG
TATCCCGACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGA
AAGAGTTCCTCGACGCTAACCTGGCGTAAGCGGCCGC

(SEQ ID NO: 203)
OST269 (irrelevant control peptide for OST266-OST268)
AAGCTTGCCACCATGGGCGATAAAATTATTCACCTGACTGACGACAGTTTTGACACGGATGTACTCAAAGCGGACGG
GGCGATCCTCGTCGATTTCTGGGCAGAGTGGTGCGGTCCGAATTCCTTGTGGATTGGGCCGGGTGATCAGGGTCTGT
TTCGGCGTTTTGTTTTTACTAGTGGTCCGTGCAAAATGATCGCCCCGATTCTGGATGAAATCGCTGACGAATATCAG
GGCAAACTGACCGTTGCAAAACTGAACATCGATCAAAACCCTGGCACTGCGCCGAAATATGGCATCCGTGGTATCCC
GACTCTGCTGCTGTTCAAAAACGGTGAAGTGGCGGCAACCAAAGTGGGTGCACTGTCTAAAGGTCAGTTGAAAGAGT
TCCTCGACGCTAACCTGGCGTAAGCGGCCGC

FIGURE 12B

Listed below are the amino acid sequences corresponding to the pcDNA3.1 constructs in Appendix 1A

OST258
METDTLLLWVLLLWVPGSTGDGS

OST259 (SEQ ID NO: 204)
METDTLLLWVLLLWVPGSTGDGSMSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPNSVVLCSRCGRLWRWSCGT
SGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA

OST260 (SEQ ID NO: 205)
METDTLLLWVLLLWVPGSTGDGSMSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPNSGWRWCGRCGALWWRRVT
SGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA

OST261 (SEQ ID NO: 206)
METDTLLLWVLLLWVPGSTGDGSMSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPNSEVRQVTCIRCRRGFLLT
SGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA

OST262 (SEQ ID NO: 207)
METDTLLLWVLLLWVPGSTGDGSMSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPNSGGGGMIWEAWSCYACGT
SGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA

OST263 (SEQ ID NO: 208)
METDTLLLWVLLLWVPGSTGDGSMSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPNSLWIGPGDQGLFRRFVFT
SGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA

OST264 (SEQ ID NO: 209)
METDTLLLWVLLLWVPGSTGDGSVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEGSVCLRS
SDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQKDHHQASNSSRLHTCQRH

OST265 (SEQ ID NO: 210)
METDTLLLWVLLLWVPGSTGDGSCASPTRGGDAGVQICLACRKRRKRCMRHAMCCPGNYCKNGICVSSDQNHFRGEI
EETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEGSVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKG
SHGLEI
FQRCYCGEGLSC.

OST266 (SEQ ID NO: 211)
MGDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPNSYAWLFSCSRCRWWLPWTSGPCKMIAPILDEIADEYQGKLT
VAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA

OST267 (SEQ ID NO: 212)
MGDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPNSICEVVRLWSRYPWSWVTSGPCKMIAPILDEIADEYQGKLT
VAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA

OST268 (SEQ ID NO: 213)
MGDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPNSGCTSAVCGAWAEAGRFYCTSGPCKMIAPILDEIADEYQGK
LTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA

OST269 (SEQ ID NO: 214)
MGDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPNSLWIGPGDQGLFRRFVFTSGPCKMIAPILDEIADEYQGKLT
VAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLA

FIGURE 13

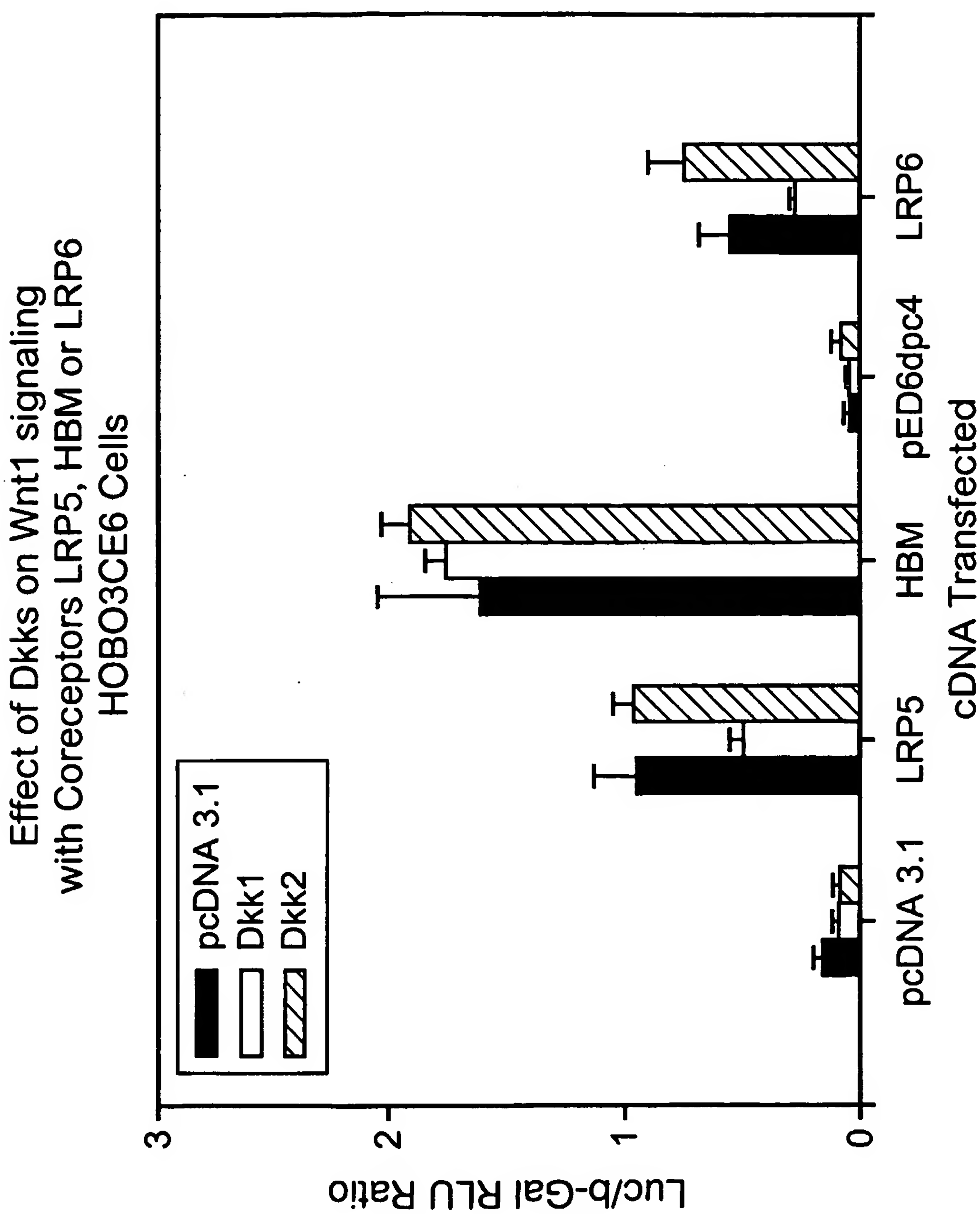

Effect of Dkks on Wnt1 signaling
with Coreceptors LRP5, HBM or LRP6
HOBO3CE6 Cells
pcDNA 3.1
Dkk1
Dkk2
Luc/b-Gal RLU Ratio
0
1
2
3
pcDNA 3.1
LRP5
HBM
pED6dpc4
LRP6
cDNA Transfected

FIG. 14

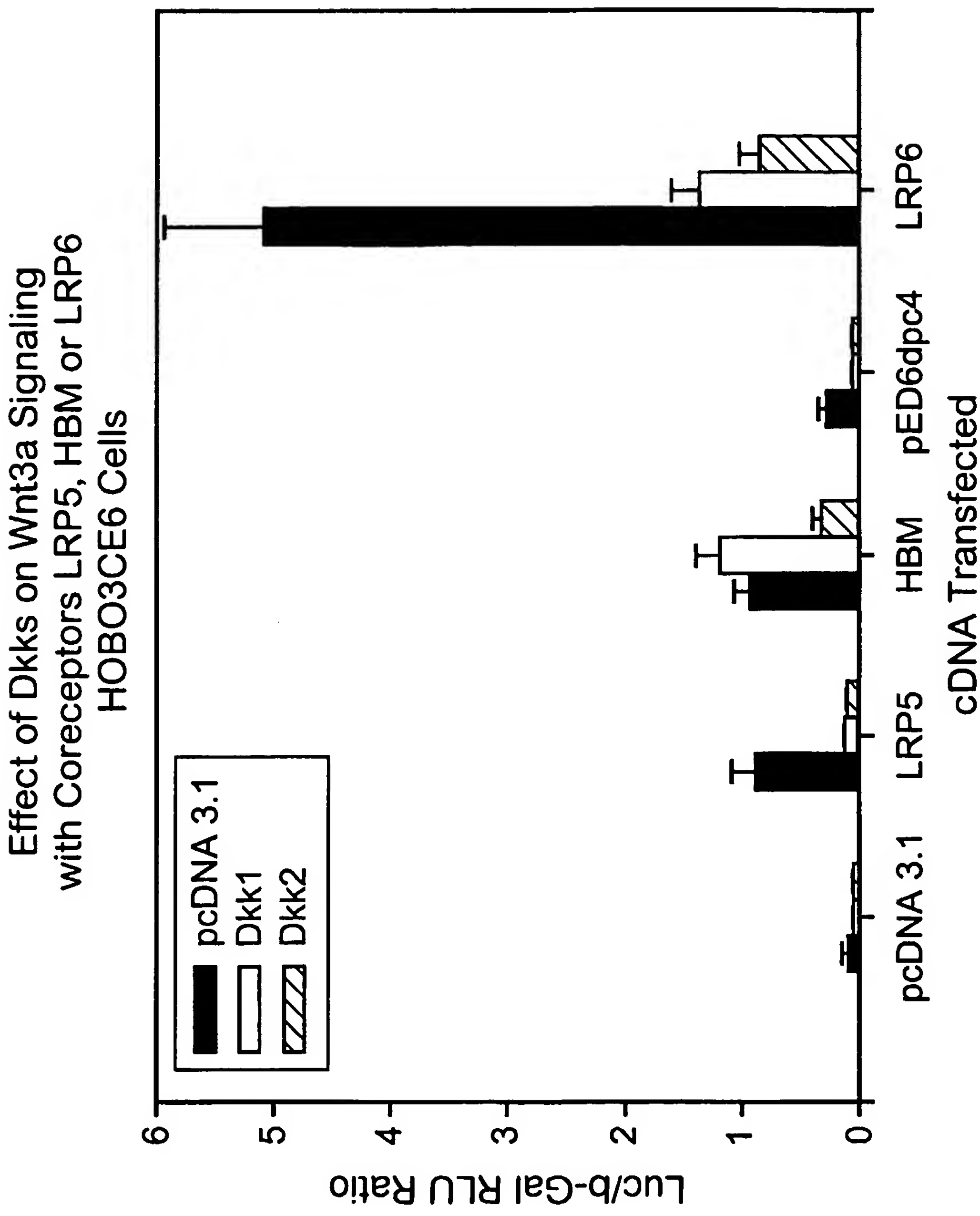

Effect of Dkks on Wnt3a Signaling
with Coreceptors LRP5, HBM or LRP6
HOBO3CE6 Cells
pcDNA 3.1
Dkk1
Dkk2
Luc/b-Gal RLU Ratio
0
1
2
3
4
5
6
pcDNA 3.1
LRP5
HBM
pED6dpc4
LRP6
cDNA Transfected

FIG. 15

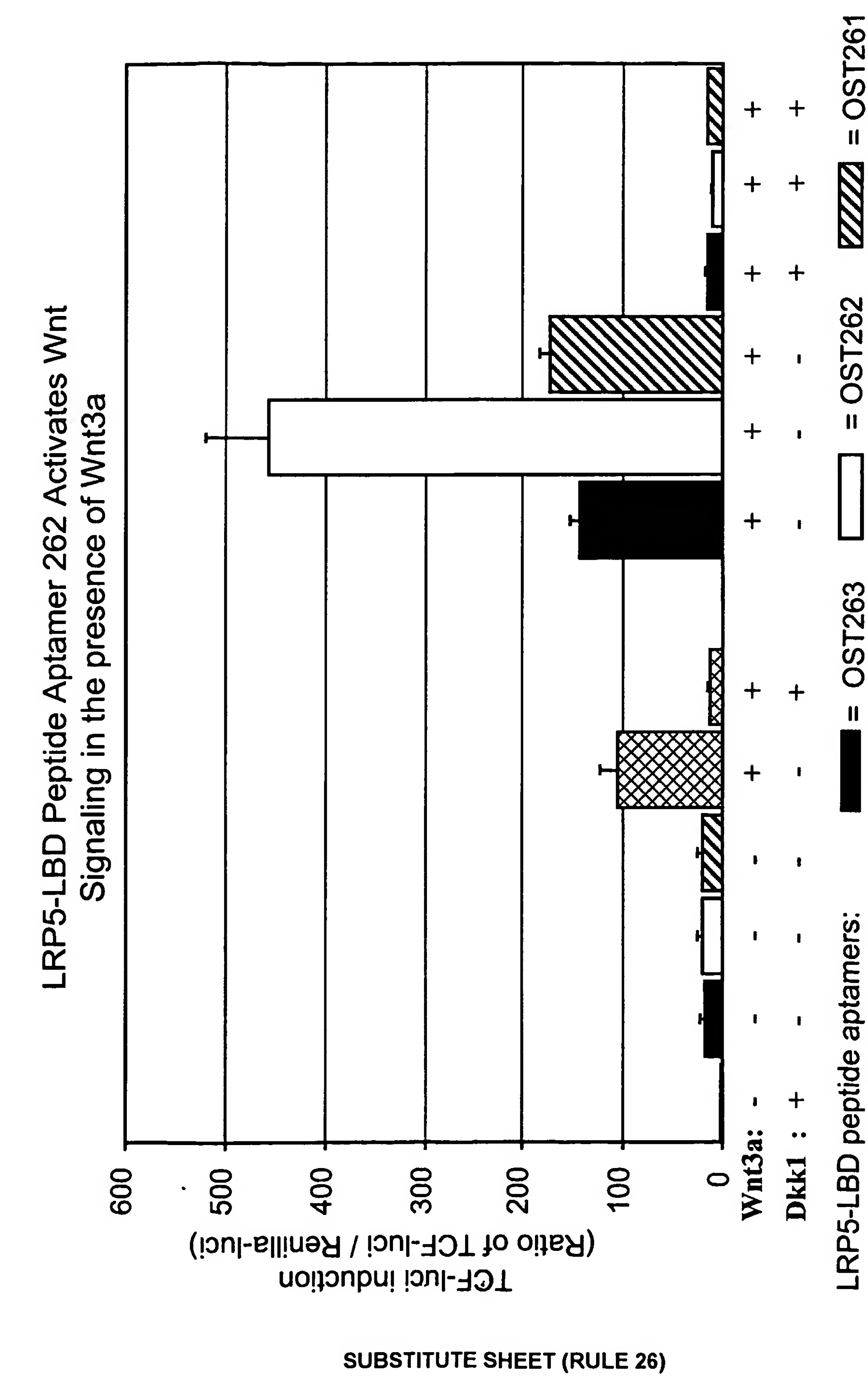
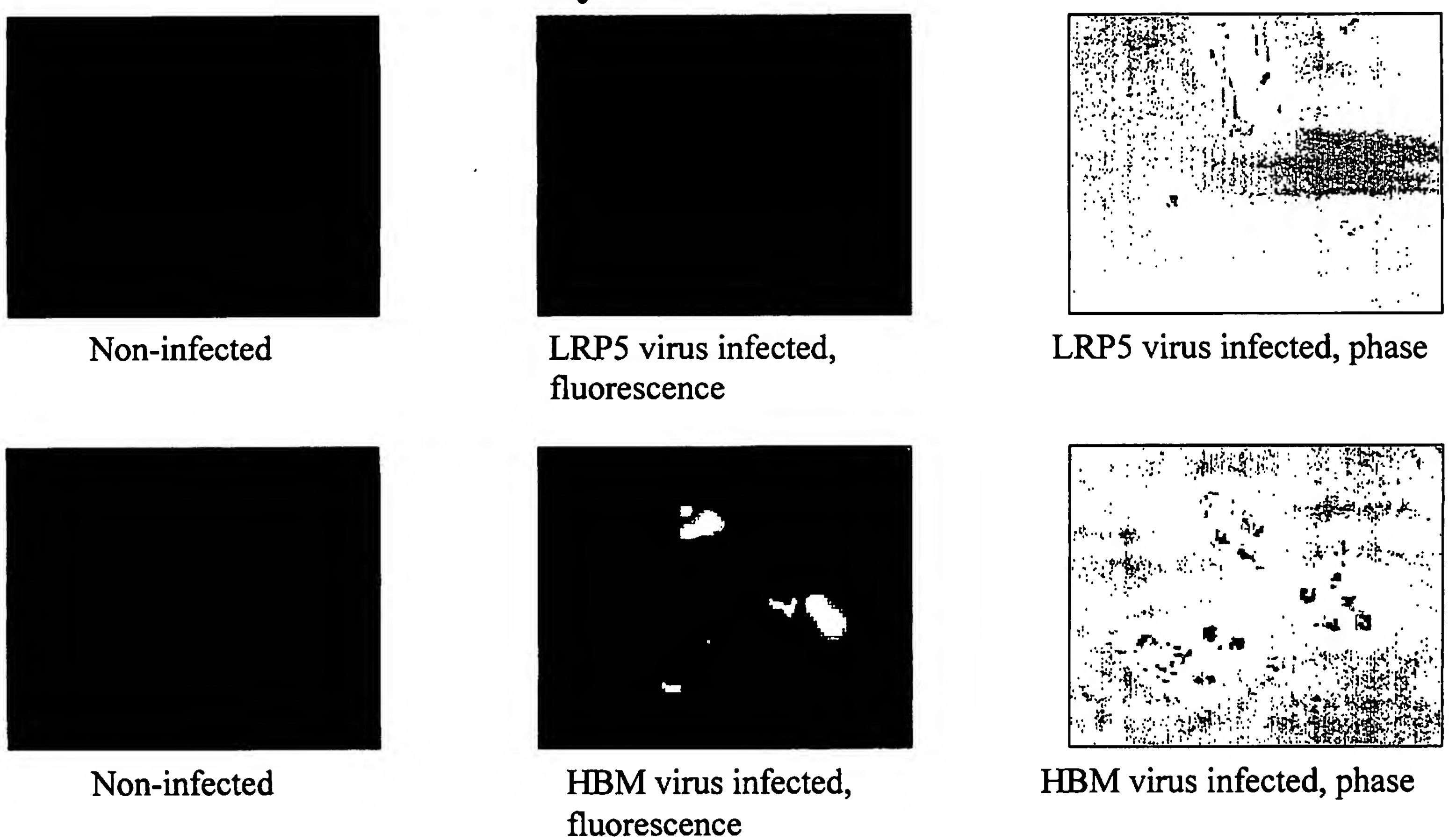

LRP5-LBD Peptide Aptamer 262 Activates Wnt Signaling in the presence of Wnt3a
TCF-luci induction (Ratio of TCF-luci / Renilla-luci)
600
500
400
300
200
100
0
Wnt3a: - - - - + + + + + + +
Dkk1 : + - - - - + - - - + + +
LRP5-LBD peptide aptamers:
= OST263
= OST262
= OST261

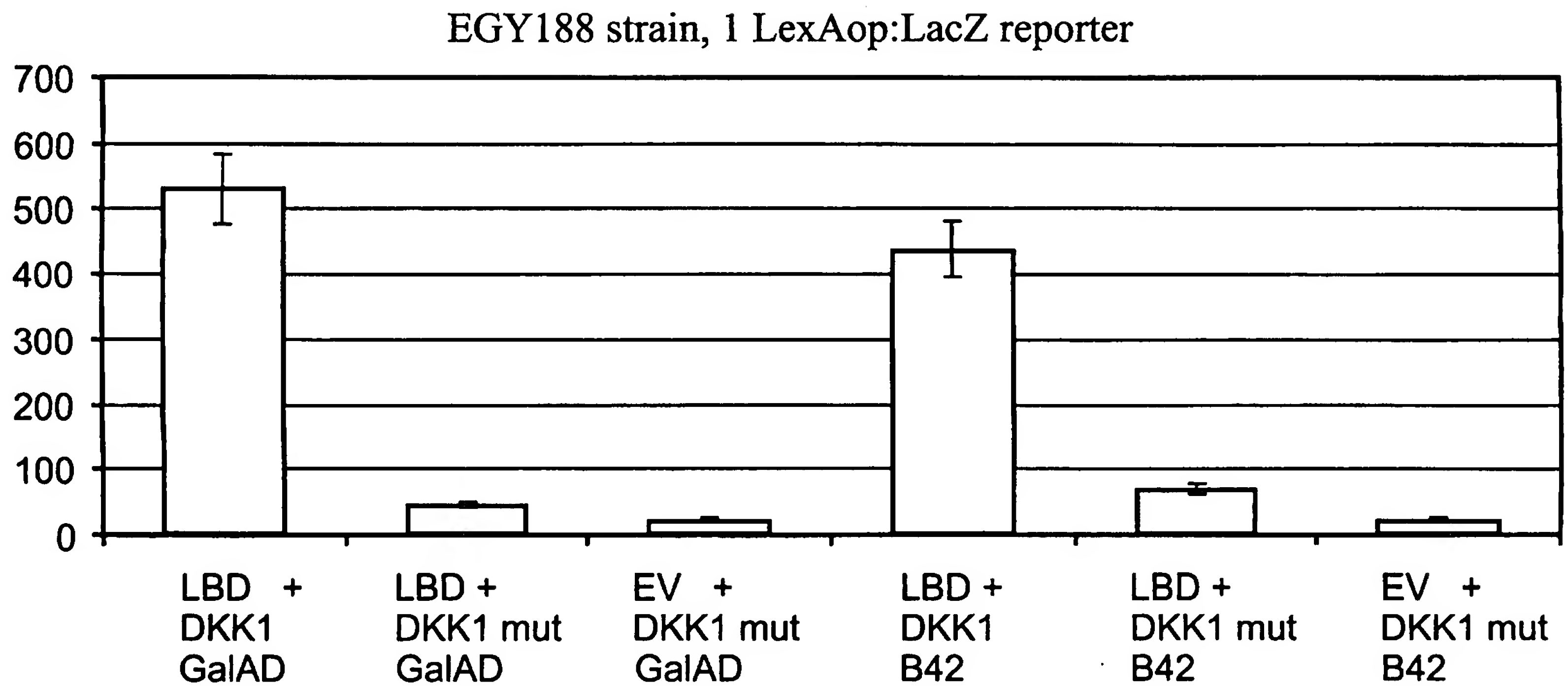
FIG. 16

FIG. 17

Affinity Purified 69546/47



Antibody to: aa 165-177 (Mutation)



- A mutant DKK1,C220A, unable to bind to LRP5, was cloned in GalAD and B42 and tested for its ability to bind to LBD in Y2H
- Interaction LBD-DKK1 20 fold above background
- Interaction LBD-DKK1 C220A 2 to 3 fold above background
- Interaction LBD-DKK1 10 fold above LBD-DKK1 C220A mutant

FIG. 18

FIG. 19

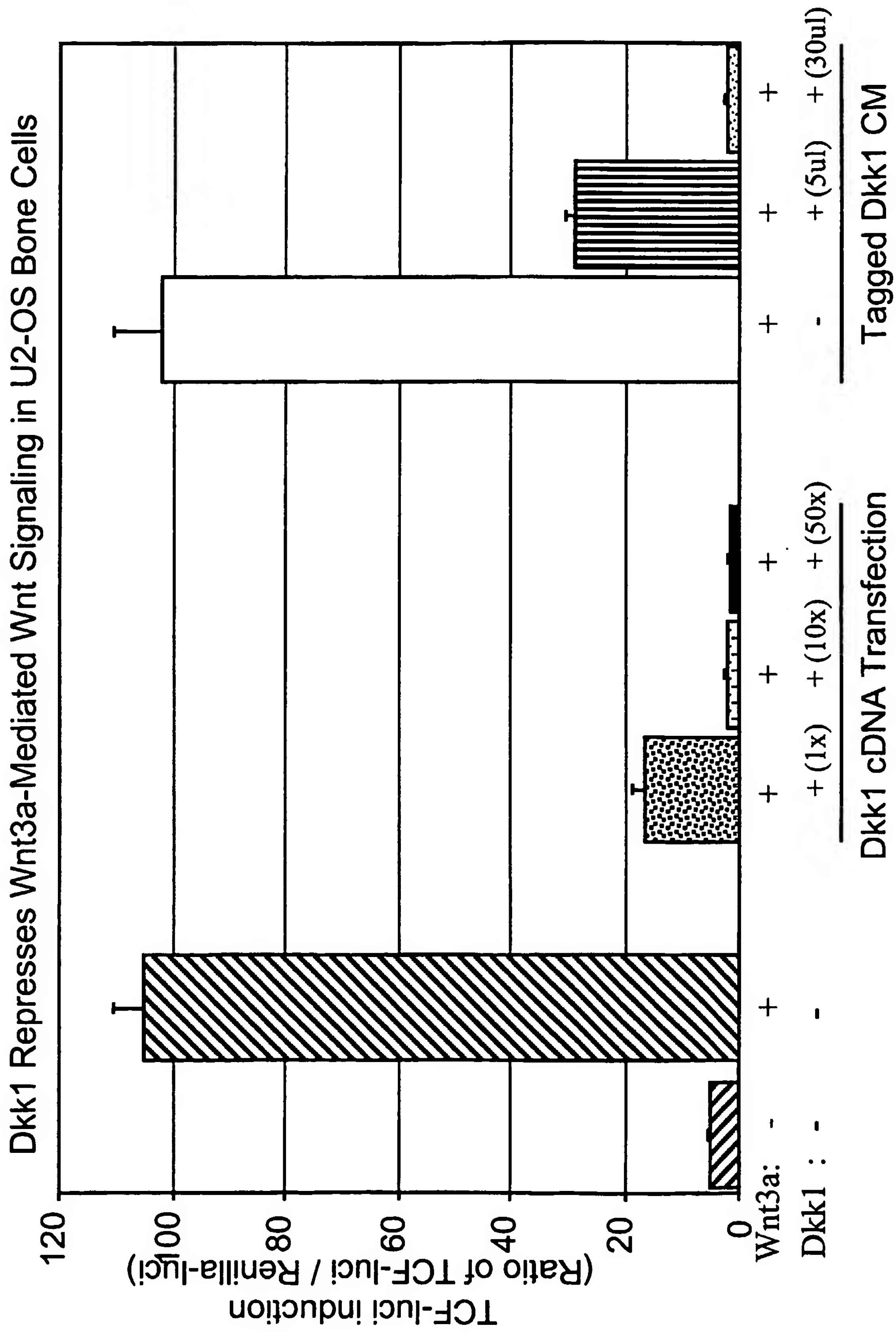
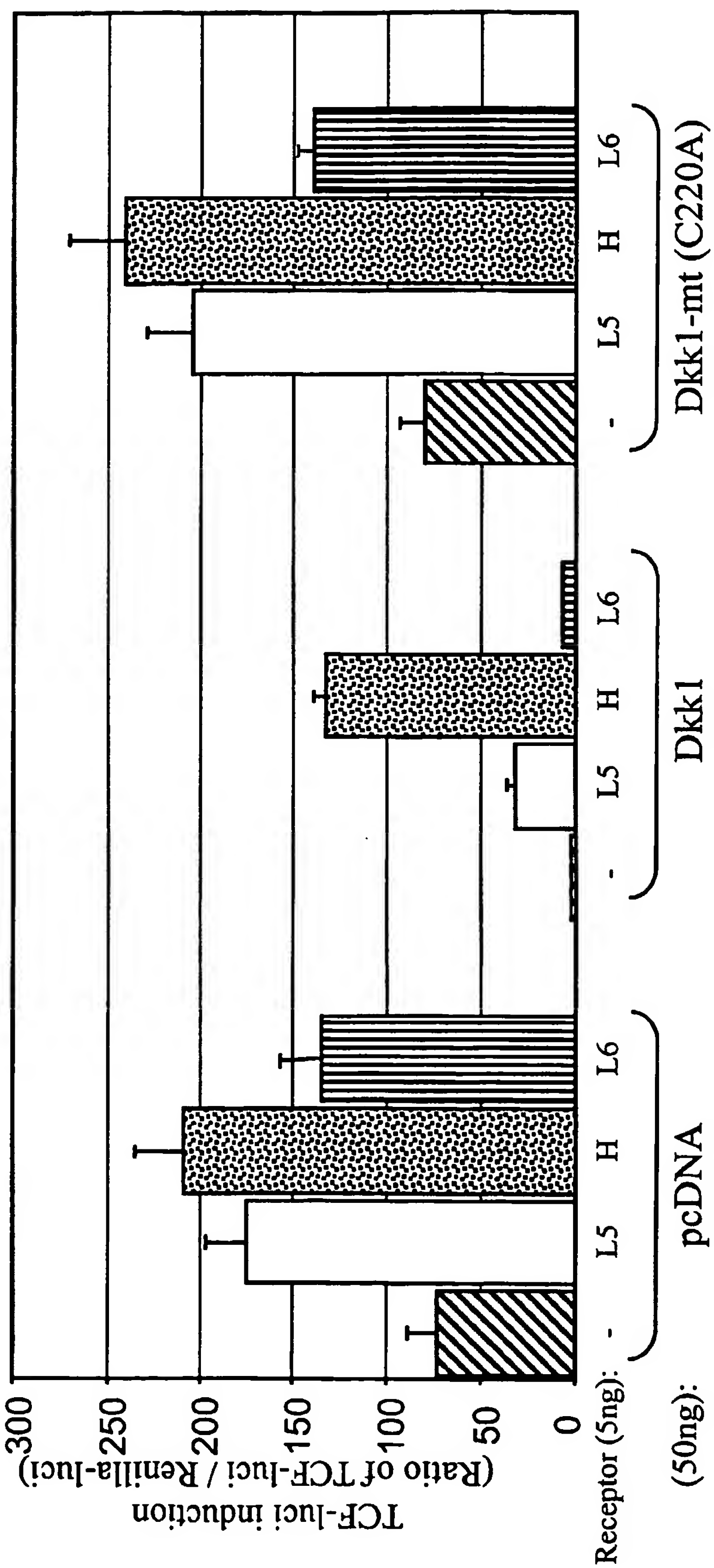

Dkk1 Represses Wnt3a-Mediated Wnt Signaling in U2-OS Bone Cells
TCF-luci induction
(Ratio of TCF-luci / Renilla-luci)
120
100
80
60
40
20
0
Wnt3a: - + + + + + + +
Dkk1 : - - + (1x) + (10x) + (50x) - + (5ul) + (30ul)
Dkk1 cDNA Transfection
Tagged Dkk1 CM

FIG. 20

Wnt1 - HBM generated TCF-luci is not efficiently inhibited by Dkk1 in U2-OS bone cells.



- With Wnt1 the TCF-signal generated by LRP5 is greater than that of LRP6.
- LRP5/6 -Wnt1 induced TCF- is efficiently blocked byDkk1

In U2-OS cells TCF-signal can be modulated by Dkk1, Dkk1-AP, without Wnt DNA transfection.

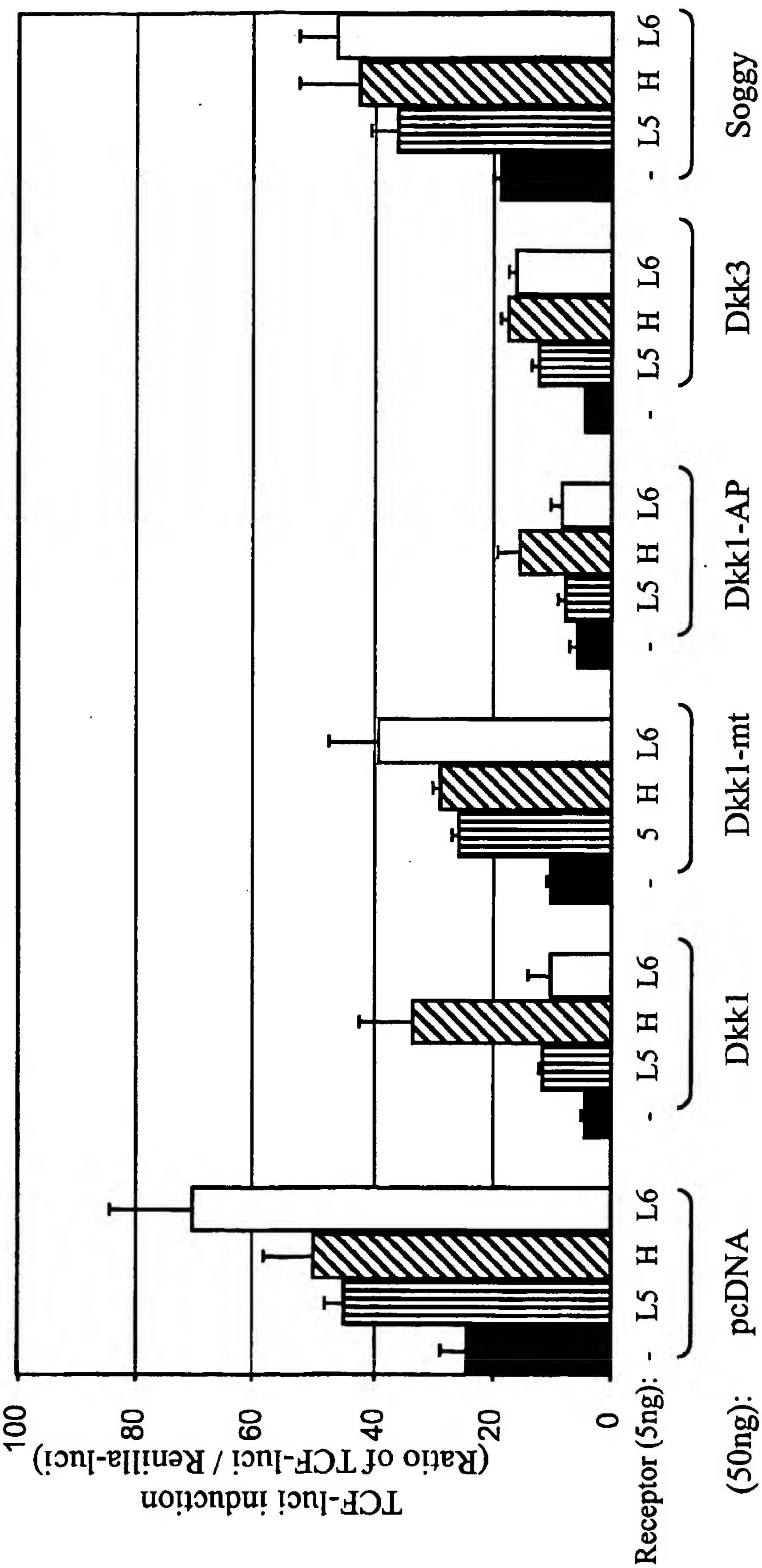


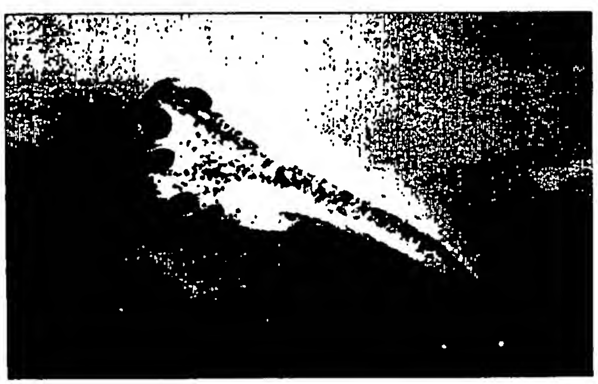
FIG. 21

FIG. 22

Aptamers 261 and 262 from the LRP5-LBD Activate Wnt Signaling in Xenopus

263 - Negative Control

261 - LBD-Binding Peptide

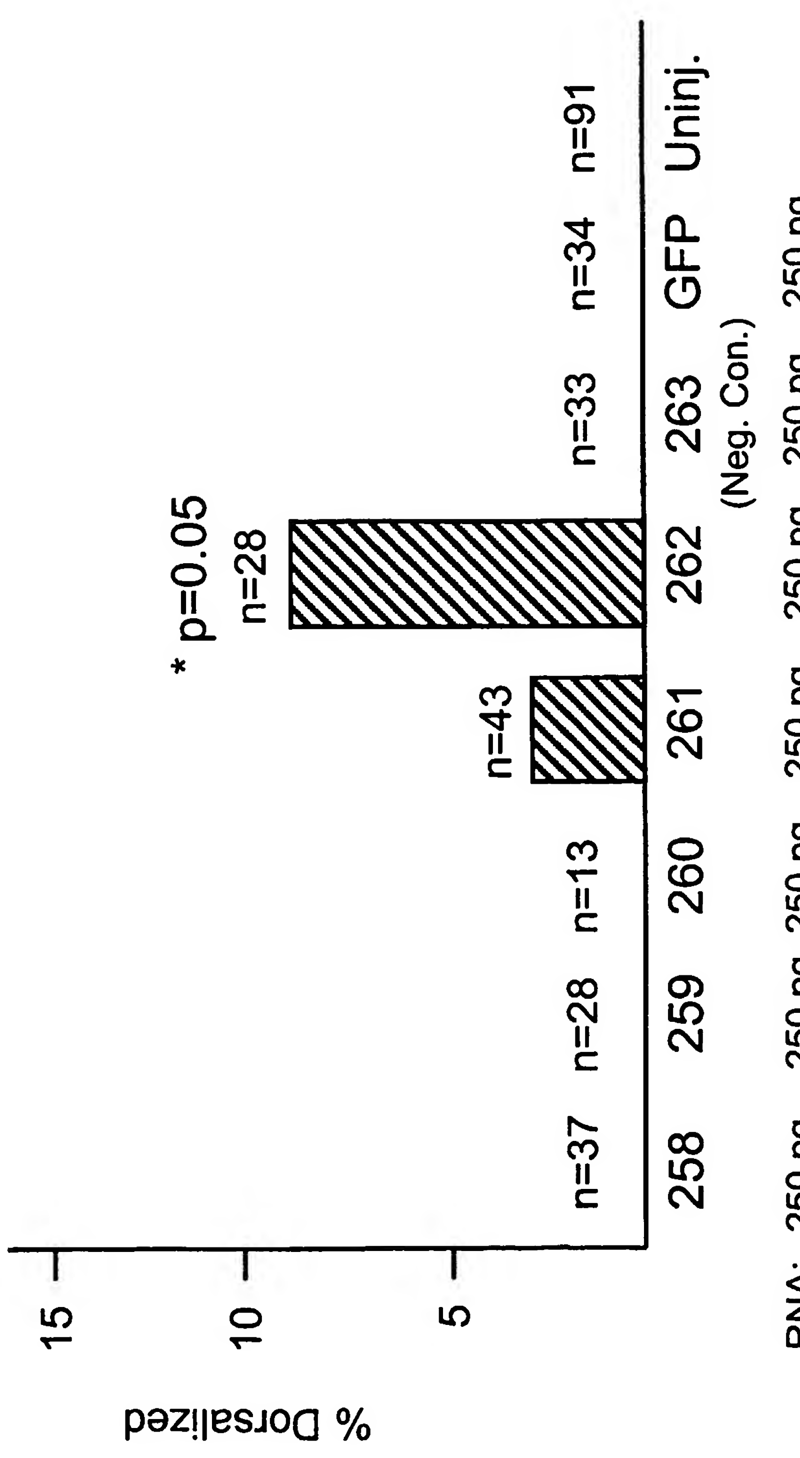


262 - LBD-Binding Peptide

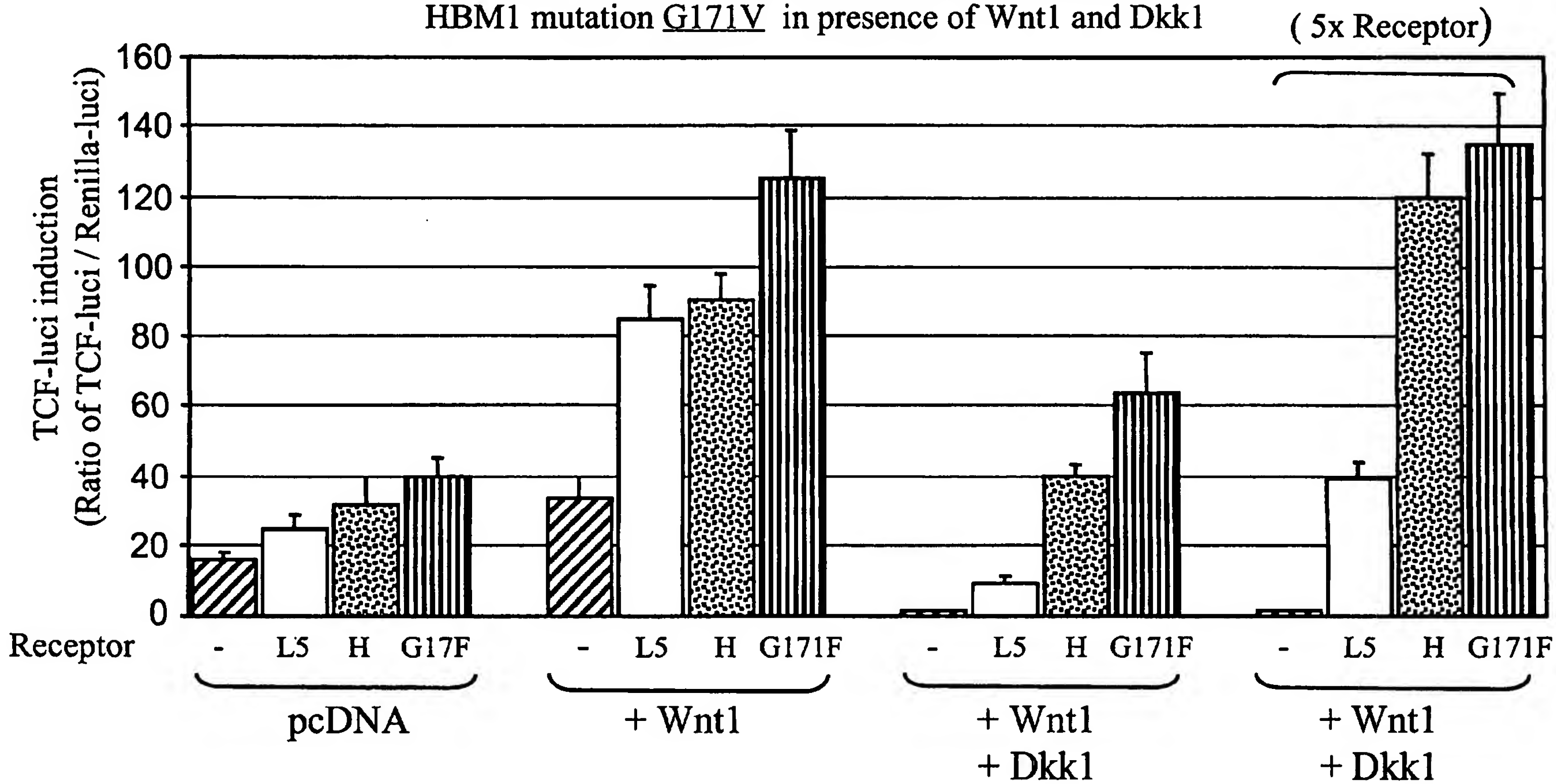
262 - LBD-Binding Peptide

FIG. 23

LRP5 Peptide Aptamers 261 and 262 Induce Wnt Signaling



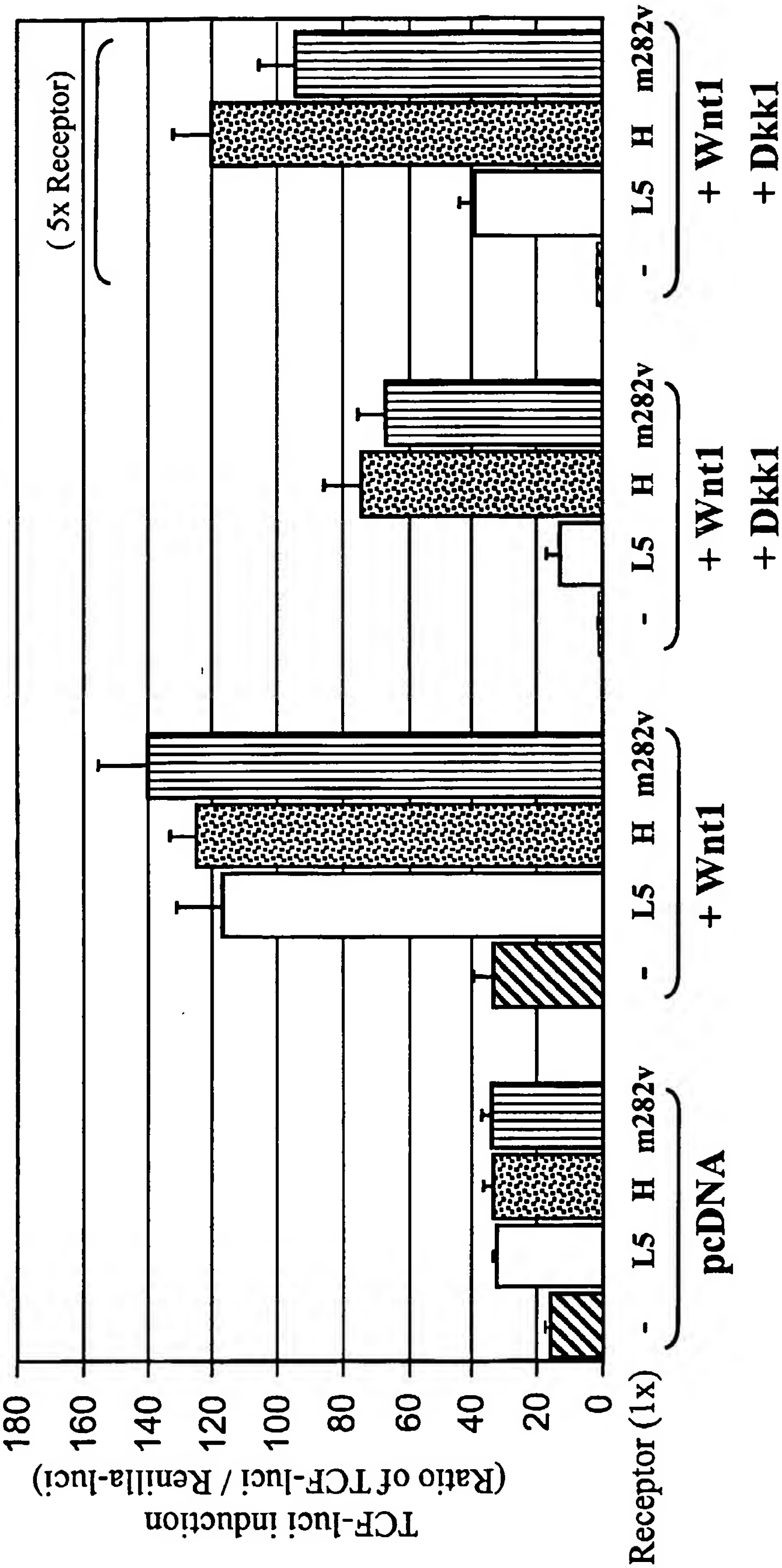
LRP 5 mutation G171F results in greater TCF-activation than with HBM1 mutation G171V in presence of Wnt1 and Dkk1



• G171F mutation involves the ringed R group (F) alteration and leads to marginally greater TCF-luci activation than that with HBM1 mutation G171V.

FIG. 24

LRP5-blade1 Mutant M282V leads to HBM1 type TCF-signal with Wnt1 and Dkk1 in U2-OS cells



- In blade 1, propeller 1, M282 is at the accessible interior position.
- It is conserved in propellers 1-3

FIG. 25

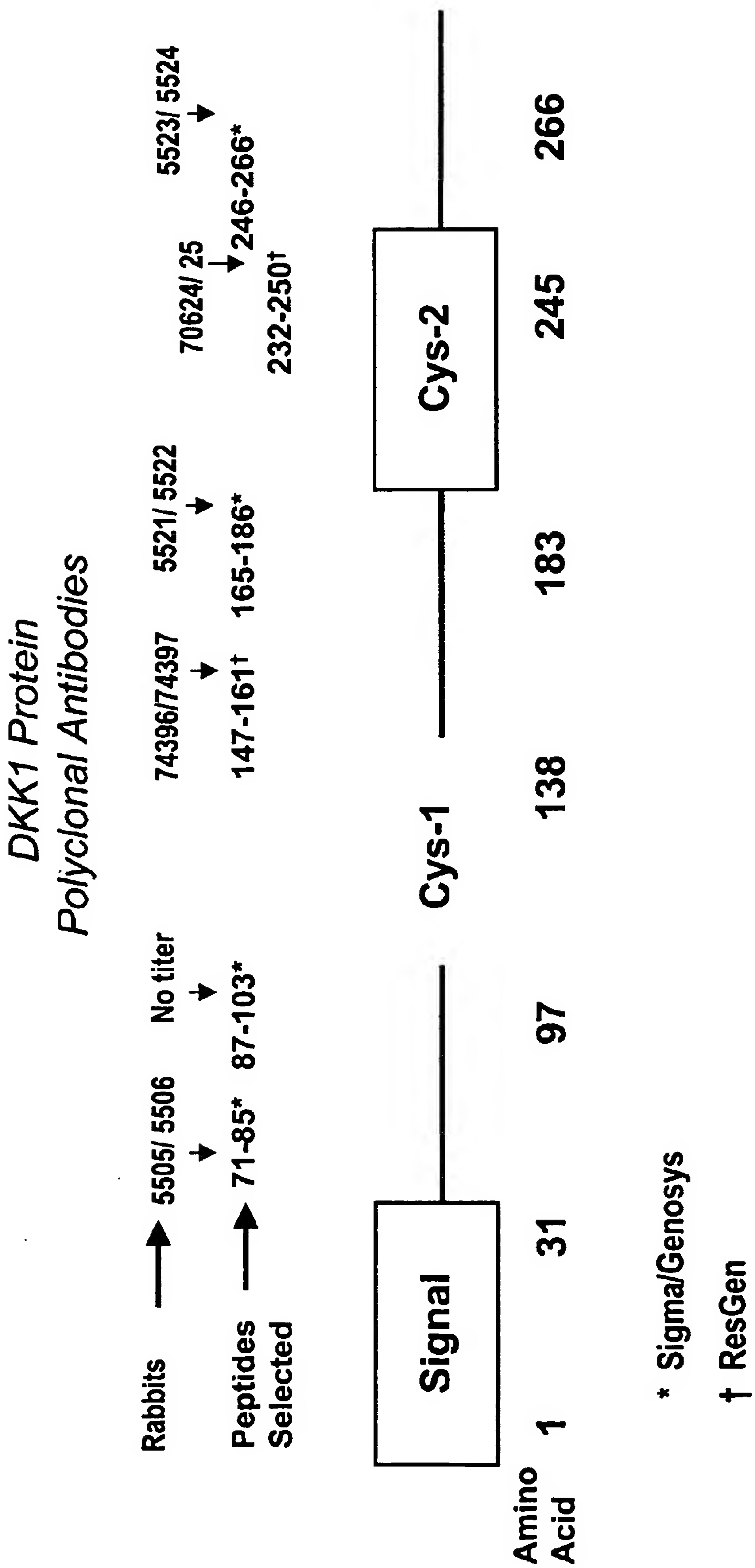

DKK1 Protein
Polyclonal Antibodies
Rabbits
Peptides Selected
5505/ 5506
71-85*
No titer
87-103*
74396/74397
147-161†
5521/ 5522
165-186*
70624/ 25
232-250†
5523/ 5524
246-266*
Signal
Cys-1
Cys-2
Amino Acid
1
31
97
138
183
245
266
* Sigma/Genosys
† ResGen

FIG. 26

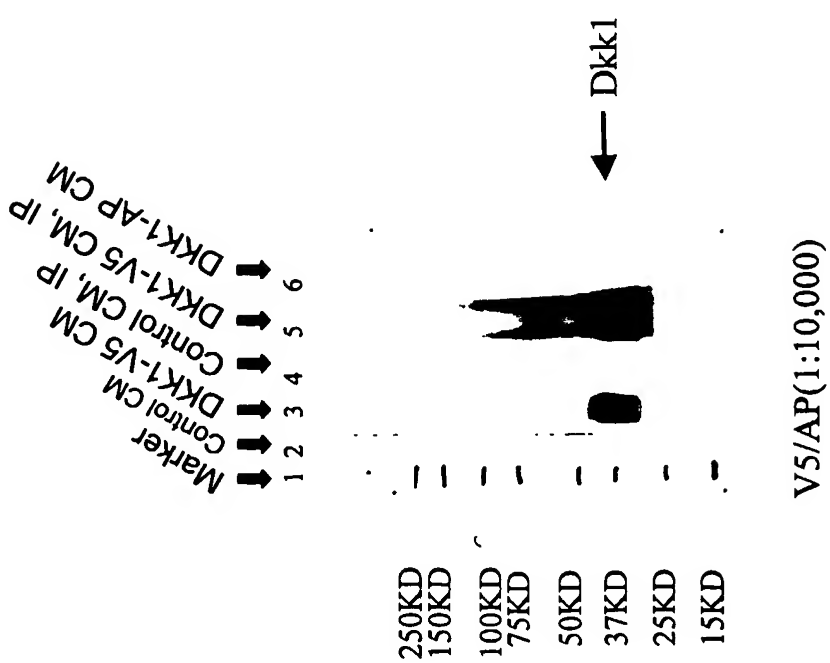


FIG. 27A

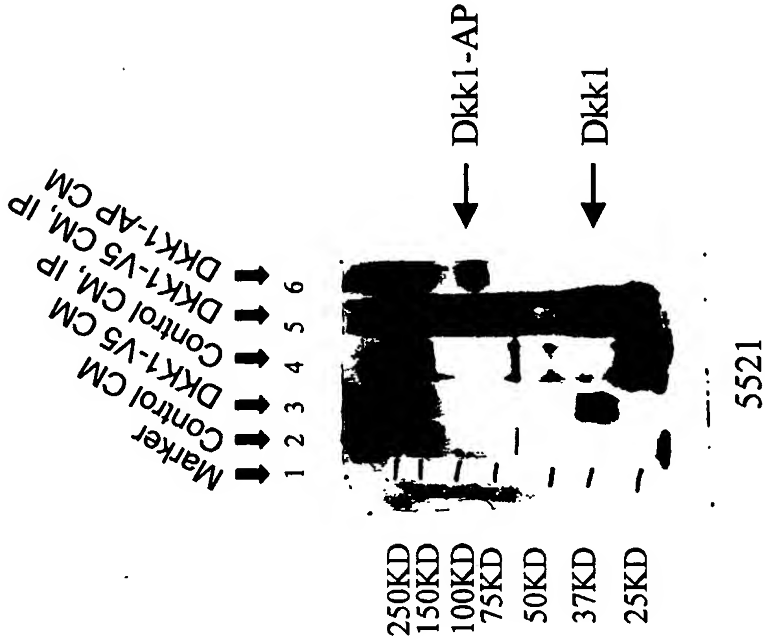


FIG. 27B

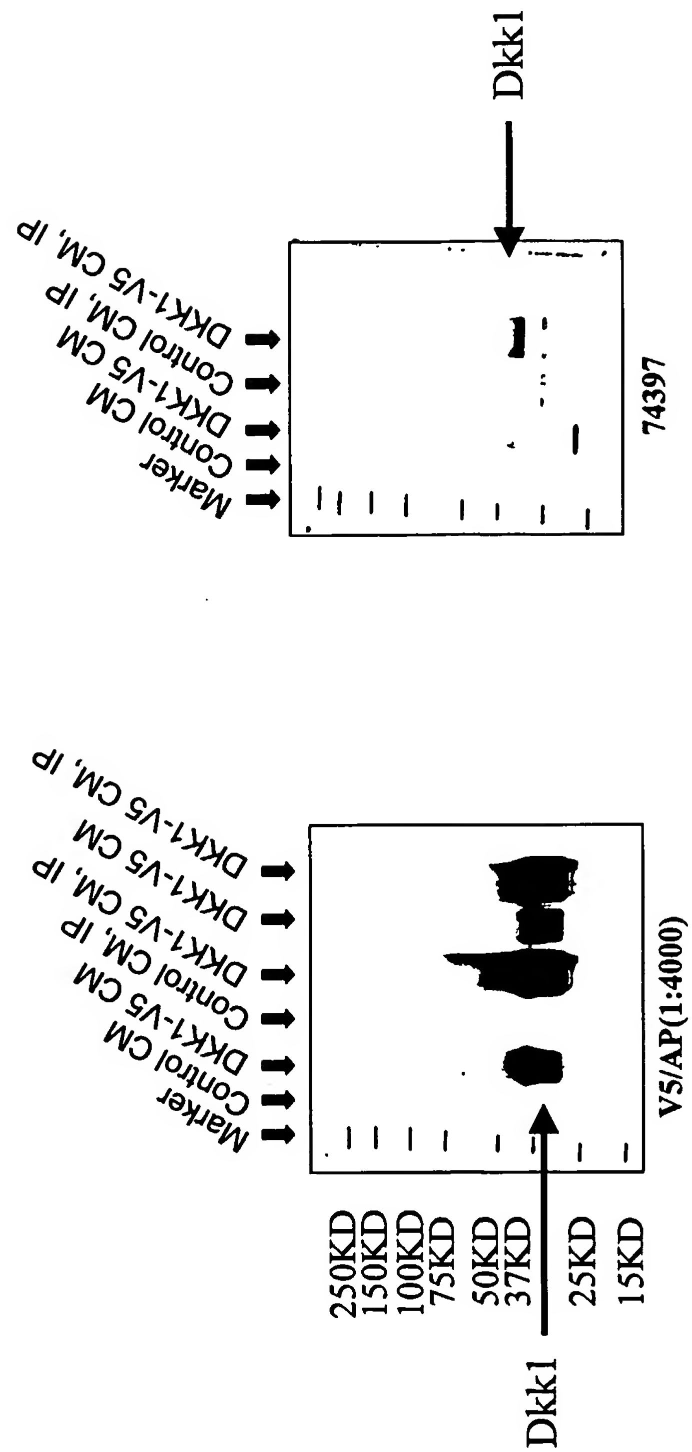

Marker
Control CM
DKK1-V5 CM
Control CM, IP
DKK1-V5 CM, IP
DKK1-V5 CM
DKK1-V5 CM, IP
250KD
150KD
100KD
75KD
50KD
37KD
25KD
15KD
Dkk1
V5/AP(1:4000)
Marker
Control CM
DKK1-V5 CM
Control CM, IP
DKK1-V5 CM, IP
Dkk1
74397

FIG. 28

032796-132.ST25

SEQUENCE LISTING

<110> Allen, Kristina M.
Anisowicz, Anthony
Bhat, Bheem
Damagnez, Veronique
Robinson, John
Yaworsky, Paul

<120> Reagents and Method for Modulating DKK-Mediated Interactions

<130> 032796-132

<150> US 60/291,311
<151> 2001-05-17

<150> US 60/353,058
<151> 2002-02-01

<150> US 60/361,293
<151> 2002-03-04

<160> 214

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 5120
<212> DNA
<213> Homo sapiens

<400> 1

```
actaaagcgc cgccgccgcg ccatggagcc cgagtgagcg cggcgcgggc ccgtccggcc        60
gccggacaac  atg gag gca gcg ccg ccc ggg ccg ccg tgg ccg ctg ctg         109
            Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu
            1               5                   10
 ctg ctg ctg ctg ctg ctg ctg gcg ctg tgc ggc tgc ccg gcc ccc gcc        157
 Leu Leu Leu Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala
     15                  20                  25
 gcg gcc tcg ccg ctc ctg cta ttt gcc aac cgc cgg gac gta cgg ctg        205
 Ala Ala Ser Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu
 30                  35                  40                  45
 gtg gac gcc ggc gga gtc aag ctg gag tcc acc atc gtg gtc agc ggc        253
 Val Asp Ala Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly
                 50                  55                  60
 ctg gag gat gcg gcc gca gtg gac ttc cag ttt tcc aag gga gcc gtg        301
 Leu Glu Asp Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val
             65                  70                  75
 tac tgg aca gac gtg agc gag gag gcc atc aag cag acc tac ctg aac        349
 Tyr Trp Thr Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn
         80                  85                  90
 cag acg ggg gcc gcc gtg cag aac gtg gtc atc tcc ggc ctg gtc tct        397
```

```
Gln Thr Gly Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser
    95                  100                 105
ccc gac ggc ctc gcc tgc gac tgg gtg ggc aag aag ctg tac tgg acg       445
Pro Asp Gly Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr
110                 115                 120                 125
gac tca gag acc aac cgc atc gag gtg gcc aac ctc aat ggc aca tcc       493
Asp Ser Glu Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser
                130                 135                 140
cgg aag gtg ctc ttc tgg cag gac ctt gac cag ccg agg gcc atc gcc       541
Arg Lys Val Leu Phe Trp Gln Asp Leu Asp Gln Pro Arg Ala Ile Ala
            145                 150                 155
ttg gac ccc gct cac ggg tac atg tac tgg aca gac tgg ggt gag acg       589
Leu Asp Pro Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Gly Glu Thr
        160                 165                 170
ccc cgg att gag cgg gca ggg atg gat ggc agc acc cgg aag atc att       637
Pro Arg Ile Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile
    175                 180                 185
gtg gac tcg gac att tac tgg ccc aat gga ctg acc atc gac ctg gag       685
Val Asp Ser Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu
190                 195                 200                 205
gag cag aag ctc tac tgg gct gac gcc aag ctc agc ttc atc cac cgt       733
Glu Gln Lys Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg
                210                 215                 220
gcc aac ctg gac ggc tcg ttc cgg cag aag gtg gtg gag ggc agc ctg       781
Ala Asn Leu Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu
            225                 230                 235
acg cac ccc ttc gcc ctg acg ctc tcc ggg gac act ctg tac tgg aca       829
Thr His Pro Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr
        240                 245                 250
gac tgg cag acc cgc tcc atc cat gcc tgc aac aag cgc act ggg ggg       877
Asp Trp Gln Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly
    255                 260                 265
aag agg aag gag atc ctg agt gcc ctc tac tca ccc atg gac atc cag       925
Lys Arg Lys Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln
270                 275                 280                 285
gtg ctg agc cag gag cgg cag cct ttc ttc cac act cgc tgt gag gag       973
Val Leu Ser Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu
                290                 295                 300
gac aat ggc ggc tgc tcc cac ctg tgc ctg ctg tcc cca agc gag cct       1021
Asp Asn Gly Gly Cys Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro
            305                 310                 315
ttc tac aca tgc gcc tgc ccc acg ggt gtg cag ctg cag gac aac ggc       1069
Phe Tyr Thr Cys Ala Cys Pro Thr Gly Val Gln Leu Gln Asp Asn Gly
        320                 325                 330
agg acg tgt aag gca gga gcc gag gag gtg ctg ctg ctg gcc cgg cgg       1117
Arg Thr Cys Lys Ala Gly Ala Glu Glu Val Leu Leu Leu Ala Arg Arg
    335                 340                 345
acg gac cta cgg agg atc tcg ctg gac acg ccg gac ttc acc gac atc       1165
Thr Asp Leu Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile
350                 355                 360                 365
gtg ctg cag gtg gac gac atc cgg cac gcc att gcc atc gac tac gac       1213
Val Leu Gln Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp
                370                 375                 380
ccg cta gag ggc tat gtc tac tgg aca gat gac gag gtg cgg gcc atc       1261
Pro Leu Glu Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile
            385                 390                 395
cgc agg gcg tac ctg gac ggg tct ggg gcg cag acg ctg gtc aac acc       1309
```

Arg Arg Ala Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr
400 405 410
gag atc aac gac ccc gat ggc atc gcg gtc gac tgg gtg gcc cga aac 1357
Glu Ile Asn Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn
415 420 425
ctc tac tgg acc gac acg ggc acg gac cgc atc gag gtg acg cgc ctc 1405
Leu Tyr Trp Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu
430 435 440 445
aac ggc acc tcc cgc aag atc ctg gtg tcg gag gac ctg gac gag ccc 1453
Asn Gly Thr Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro
450 455 460
cga gcc atc gca ctg cac ccc gtg atg ggc ctc atg tac tgg aca gac 1501
Arg Ala Ile Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp
465 470 475
tgg gga gag aac cct aaa atc gag tgt gcc aac ttg gat ggg cag gag 1549
Trp Gly Glu Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu
480 485 490
cgg cgt gtg ctg gtc aat gcc tcc ctc ggg tgg ccc aac ggc ctg gcc 1597
Arg Arg Val Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala
495 500 505
ctg gac ctg cag gag ggg aag ctc tac tgg gga gac gcc aag aca gac 1645
Leu Asp Leu Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp
510 515 520 525
aag atc gag gtg atc aat gtt gat ggg acg aag agg cgg acc ctc ctg 1693
Lys Ile Glu Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu
530 535 540
gag gac aag ctc ccg cac att ttc ggg ttc acg ctg ctg ggg gac ttc 1741
Glu Asp Lys Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe
545 550 555
atc tac tgg act gac tgg cag cgc cgc agc atc gag cgg gtg cac aag 1789
Ile Tyr Trp Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys
560 565 570
gtc aag gcc agc cgg gac gtc atc att gac cag ctg ccc gac ctg atg 1837
Val Lys Ala Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met
575 580 585
ggg ctc aaa gct gtg aat gtg gcc aag gtc gtc gga acc aac ccg tgt 1885
Gly Leu Lys Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys
590 595 600 605
gcg gac agg aac ggg ggg tgc agc cac ctg tgc ttc ttc aca ccc cac 1933
Ala Asp Arg Asn Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro His
610 615 620
gca acc cgg tgt ggc tgc ccc atc ggc ctg gag ctg ctg agt gac atg 1981
Ala Thr Arg Cys Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met
625 630 635
aag acc tgc atc gtg cct gag gcc ttc ttg gtc ttc acc agc aga gcc 2029
Lys Thr Cys Ile Val Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala
640 645 650
gcc atc cac agg atc tcc ctc gag acc aat aac aac gac gtg gcc atc 2077
Ala Ile His Arg Ile Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile
655 660 665
ccg ctc acg ggc gtc aag gag gcc tca gcc ctg gac ttt gat gtg tcc 2125
Pro Leu Thr Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser
670 675 680 685
aac aac cac atc tac tgg aca gac gtc agc ctg aag acc atc agc cgc 2173
Asn Asn His Ile Tyr Trp Thr Asp Val Ser Leu Lys Thr Ile Ser Arg
690 695 700
gcc ttc atg aac ggg agc tcg gtg gag cac gtg gtg gag ttt ggc ctt 2221

```
Ala Phe Met Asn Gly Ser Ser Val Glu His Val Val Glu Phe Gly Leu
            705                 710                 715
gac tac ccc gag ggc atg gcc gtt gac tgg atg ggc aag aac ctc tac        2269
Asp Tyr Pro Glu Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr
        720                 725                 730
tgg gcc gac act ggg acc aac aga atc gaa gtg gcg cgg ctg gac ggg        2317
Trp Ala Asp Thr Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly
    735                 740                 745
cag ttc cgg caa gtc ctc gtg tgg agg gac ttg gac aac ccg agg tcg        2365
Gln Phe Arg Gln Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser
750                 755                 760                 765
ctg gcc ctg gat ccc acc aag ggc tac atc tac tgg acc gag tgg ggc        2413
Leu Ala Leu Asp Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly
                770                 775                 780
ggc aag ccg agg atc gtg cgg gcc ttc atg gac ggg acc aac tgc atg        2461
Gly Lys Pro Arg Ile Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met
            785                 790                 795
acg ctg gtg gac aag gtg ggc cgg gcc aac gac ctc acc att gac tac        2509
Thr Leu Val Asp Lys Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr
        800                 805                 810
gct gac cag cgc ctc tac tgg acc gac ctg gac acc aac atg atc gag        2557
Ala Asp Gln Arg Leu Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu
    815                 820                 825
tcg tcc aac atg ctg ggt cag gag cgg gtc gtg att gcc gac gat ctc        2605
Ser Ser Asn Met Leu Gly Gln Glu Arg Val Val Ile Ala Asp Asp Leu
830                 835                 840                 845
ccg cac ccg ttc ggt ctg acg cag tac agc gat tat atc tac tgg aca        2653
Pro His Pro Phe Gly Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr
                850                 855                 860
gac tgg aat ctg cac agc att gag cgg gcc gac aag act agc ggc cgg        2701
Asp Trp Asn Leu His Ser Ile Glu Arg Ala Asp Lys Thr Ser Gly Arg
            865                 870                 875
aac cgc acc ctc atc cag ggc cac ctg gac ttc gtg atg gac atc ctg        2749
Asn Arg Thr Leu Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu
        880                 885                 890
gtg ttc cac tcc tcc cgc cag gat ggc ctc aat gac tgt atg cac aac        2797
Val Phe His Ser Ser Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn
    895                 900                 905
aac ggg cag tgt ggg cag ctg tgc ctt gcc atc ccc ggc ggc cac cgc        2845
Asn Gly Gln Cys Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg
910                 915                 920                 925
tgc ggc tgc gcc tca cac tac acc ctg gac ccc agc agc cgc aac tgc        2893
Cys Gly Cys Ala Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys
                930                 935                 940
agc ccg ccc acc acc ttc ttg ctg ttc agc cag aaa tct gcc atc agt        2941
Ser Pro Pro Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser
            945                 950                 955
cgg atg atc ccg gac gac cag cac agc ccg gat ctc atc ctg ccc ctg        2989
Arg Met Ile Pro Asp Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu
        960                 965                 970
cat gga ctg agg aac gtc aaa gcc atc gac tat gac cca ctg gac aag        3037
His Gly Leu Arg Asn Val Lys Ala Ile Asp Tyr Asp Pro Leu Asp Lys
    975                 980                 985
ttc atc tac tgg gtg gat ggg cgc cag aac atc aag cga gcc aag gac        3085
Phe Ile Tyr Trp Val Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp
990                 995                 1000                1005
gac ggg acc cag ccc ttt gtt ttg acc tct ctg agc caa ggc caa aac        3133
```

Asp Gly Thr Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn
1010 1015 1020
cca gac agg cag ccc cac gac ctc agc atc gac atc tac agc cgg aca 3181
Pro Asp Arg Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr
1025 1030 1035
ctg ttc tgg acg tgc gag gcc acc aat acc atc aac gtc cac agg ctg 3229
Leu Phe Trp Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu
1040 1045 1050
agc ggg gaa gcc atg ggg gtg gtg ctg cgt ggg gac cgc gac aag ccc 3277
Ser Gly Glu Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro
1055 1060 1065
agg gcc atc gtc gtc aac gcg gag cga ggg tac ctg tac ttc acc aac 3325
Arg Ala Ile Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn
1070 1075 1080 1085
atg cag gac cgg gca gcc aag atc gaa cgc gca gcc ctg gac ggc acc 3373
Met Gln Asp Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr
1090 1095 1100
gag cgc gag gtc ctc ttc acc acc ggc ctc atc cgc cct gtg gcc ctg 3421
Glu Arg Glu Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu
1105 1110 1115
gtg gtg gac aac aca ctg ggc aag ctg ttc tgg gtg gac gcg gac ctg 3469
Val Val Asp Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu
1120 1125 1130
aag cgc att gag agc tgt gac ctg tca ggg gcc aac cgc ctg acc ctg 3517
Lys Arg Ile Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu
1135 1140 1145
gag gac gcc aac atc gtg cag cct ctg ggc ctg acc atc ctt ggc aag 3565
Glu Asp Ala Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys
1150 1155 1160 1165
cat ctc tac tgg atc gac cgc cag cag cag atg atc gag cgt gtg gag 3613
His Leu Tyr Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu
1170 1175 1180
aag acc acc ggg gac aag cgg act cgc atc cag ggc cgt gtc gcc cac 3661
Lys Thr Thr Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His
1185 1190 1195
ctc act ggc atc cat gca gtg gag gaa gtc agc ctg gag gag ttc tca 3709
Leu Thr Gly Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser
1200 1205 1210
gcc cac cca tgt gcc cgt gac aat ggt ggc tgc tcc cac atc tgt att 3757
Ala His Pro Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile
1215 1220 1225
gcc aag ggt gat ggg aca cca cgg tgc tca tgc cca gtc cac ctc gtg 3805
Ala Lys Gly Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val
1230 1235 1240 1245
ctc ctg cag aac ctg ctg acc tgt gga gag ccg ccc acc tgc tcc ccg 3853
Leu Leu Gln Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro
1250 1255 1260
gac cag ttt gca tgt gcc aca ggg gag atc gac tgt atc ccc ggg gcc 3901
Asp Gln Phe Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala
1265 1270 1275
tgg cgc tgt gac ggc ttt ccc gag tgc gat gac cag agc gac gag gag 3949
Trp Arg Cys Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu
1280 1285 1290
ggc tgc ccc gtg tgc tcc gcc gcc cag ttc ccc tgc gcg cgg ggt cag 3997
Gly Cys Pro Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln
1295 1300 1305
tgt gtg gac ctg cgc ctg cgc tgc gac ggc gag gca gac tgt cag gac 4045

```
Cys Val Asp Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp
1310                1315                1320                1325
cgc tca gac gag gtg gac tgt gac gcc atc tgc ctg ccc aac cag ttc        4093
Arg Ser Asp Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe
                1330                1335                1340
cgg tgt gcg agc ggc cag tgt gtc ctc atc aaa cag cag tgc gac tcc        4141
Arg Cys Ala Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser
            1345                1350                1355
ttc ccc gac tgt atc gac ggc tcc gac gag ctc atg tgt gaa atc acc        4189
Phe Pro Asp Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr
        1360                1365                1370
aag ccg ccc tca gac gac agc ccg gcc cac agc agt gcc atc ggg ccc        4237
Lys Pro Pro Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro
    1375                1380                1385
gtc att ggc atc atc ctc tct ctc ttc gtc atg ggt ggt gtc tat ttt        4285
Val Ile Gly Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe
1390                1395                1400                1405
gtg tgc cag cgc gtg gtg tgc cag cgc tat gcg ggg gcc aac ggg ccc        4333
Val Cys Gln Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro
                1410                1415                1420
ttc ccg cac gag tat gtc agc ggg acc ccg cac gtg ccc ctc aat ttc        4381
Phe Pro His Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe
            1425                1430                1435
ata gcc ccg ggc ggt tcc cag cat ggc ccc ttc aca ggc atc gca tgc        4429
Ile Ala Pro Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys
        1440                1445                1450
gga aag tcc atg atg agc tcc gtg agc ctg atg ggg ggc cgg ggc ggg        4477
Gly Lys Ser Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly
    1455                1460                1465
gtg ccc ctc tac gac cgg aac cac gtc aca ggg gcc tcg tcc agc agc        4525
Val Pro Leu Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser
1470                1475                1480                1485
tcg tcc agc acg aag gcc acg ctg tac ccg ccg atc ctg aac ccg ccg        4573
Ser Ser Ser Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro
                1490                1495                1500
ccc tcc ccg gcc acg gac ccc tcc ctg tac aac atg gac atg ttc tac        4621
Pro Ser Pro Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr
            1505                1510                1515
tct tca aac att ccg gcc act gcg aga ccg tac agg ccc tac atc att        4669
Ser Ser Asn Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile
        1520                1525                1530
cga gga atg gcg ccc ccg acg acg ccc tgc agc acc gac gtg tgt gac        4717
Arg Gly Met Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp
    1535                1540                1545
agc gac tac agc gcc agc cgc tgg aag gcc agc aag tac tac ctg gat        4765
Ser Asp Tyr Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp
1550                1555                1560                1565
ttg aac tcg gac tca gac ccc tat cca ccc cca ccc acg ccc cac agc        4813
Leu Asn Ser Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser
                1570                1575                1580
cag tac ctg tcg gcg gag gac agc tgc ccg ccc tcg ccc gcc acc gag        4861
Gln Tyr Leu Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu
            1585                1590                1595
agg agc tac ttc cat ctc ttc ccg ccc cct ccg tcc ccc tgc acg gac        4909
Arg Ser Tyr Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp
        1600                1605                1610
tca tcc tgacctcggc cgggccactc tggcttctct gtgcccctgt aaatagtttt        4965
```

```
 Ser Ser
     1615
aaatatgaac aaagaaaaaa atatatttta tgatttaaaa aataaatata attgggattt      5025
taaaaacatg agaaatgtga actgtgatgg ggtgggcagg gctgggagaa ctttgtacag      5085
tggagaaata tttataaact taattttgta aaaca                                 5120


<210> 2
<211> 5120
<212> DNA
<213> Homo sapiens


<400> 2
actaaagcgc cgccgccgcg ccatggagcc cgagtgagcg cggcgcgggc ccgtccggcc       60
gccggacaac  atg gag gca gcg ccg ccc ggg ccg ccg tgg ccg ctg ctg        109
            Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu
            1               5                   10
 ctg ctg ctg ctg ctg ctg ctg gcg ctg tgc ggc tgc ccg gcc ccc gcc       157
 Leu Leu Leu Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala
     15                  20                  25
 gcg gcc tcg ccg ctc ctg cta ttt gcc aac cgc cgg gac gta cgg ctg       205

 Ala Ala Ser Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu
 30                  35                  40                  45
 gtg gac gcc ggc gga gtc aag ctg gag tcc acc atc gtg gtc agc ggc       253
 Val Asp Ala Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly
                 50                  55                  60
 ctg gag gat gcg gcc gca gtg gac ttc cag ttt tcc aag gga gcc gtg       301
 Leu Glu Asp Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val
             65                  70                  75
 tac tgg aca gac gtg agc gag gag gcc atc aag cag acc tac ctg aac       349
 Tyr Trp Thr Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn
         80                  85                  90
 cag acg ggg gcc gcc gtg cag aac gtg gtc atc tcc ggc ctg gtc tct       397
 Gln Thr Gly Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser
     95                  100                 105
 ccc gac ggc ctc gcc tgc gac tgg gtg ggc aag aag ctg tac tgg acg       445
 Pro Asp Gly Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr
 110                 115                 120                 125
 gac tca gag acc aac cgc atc gag gtg gcc aac ctc aat ggc aca tcc       493
 Asp Ser Glu Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser
                 130                 135                 140
 cgg aag gtg ctc ttc tgg cag gac ctt gac cag ccg agg gcc atc gcc       541
 Arg Lys Val Leu Phe Trp Gln Asp Leu Asp Gln Pro Arg Ala Ile Ala
             145                 150                 155
 ttg gac ccc gct cac ggg tac atg tac tgg aca gac tgg gtt gag acg       589
 Leu Asp Pro Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Val Glu Thr
         160                 165                 170
 ccc cgg att gag cgg gca ggg atg gat ggc agc acc cgg aag atc att       637
 Pro Arg Ile Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile
     175                 180                 185
 gtg gac tcg gac att tac tgg ccc aat gga ctg acc atc gac ctg gag       685
 Val Asp Ser Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu
 190                 195                 200                 205
 gag cag aag ctc tac tgg gct gac gcc aag ctc agc ttc atc cac cgt       733
 Glu Gln Lys Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg
```

```
                210                     215                     220
gcc aac ctg gac ggc tcg ttc cgg cag aag gtg gtg gag ggc agc ctg          781
Ala Asn Leu Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu

            225                     230                     235
acg cac ccc ttc gcc ctg acg ctc tcc ggg gac act ctg tac tgg aca          829
Thr His Pro Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr
        240                     245                     250
gac tgg cag acc cgc tcc atc cat gcc tgc aac aag cgc act ggg ggg          877
Asp Trp Gln Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly
    255                     260                     265
aag agg aag gag atc ctg agt gcc ctc tac tca ccc atg gac atc cag          925
Lys Arg Lys Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln
270                 275                     280                     285
gtg ctg agc cag gag cgg cag cct ttc ttc cac act cgc tgt gag gag          973
Val Leu Ser Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu
                290                     295                     300
gac aat ggc ggc tgc tcc cac ctg tgc ctg ctg tcc cca agc gag cct          1021
Asp Asn Gly Gly Cys Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro
            305                     310                     315
ttc tac aca tgc gcc tgc ccc acg ggt gtg cag ctg cag gac aac ggc          1069
Phe Tyr Thr Cys Ala Cys Pro Thr Gly Val Gln Leu Gln Asp Asn Gly
        320                     325                     330
agg acg tgt aag gca gga gcc gag gag gtg ctg ctg ctg gcc cgg cgg          1117
Arg Thr Cys Lys Ala Gly Ala Glu Glu Val Leu Leu Leu Ala Arg Arg
    335                     340                     345
acg gac cta cgg agg atc tcg ctg gac acg ccg gac ttc acc gac atc          1165
Thr Asp Leu Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile
350                 355                     360                     365
gtg ctg cag gtg gac gac atc cgg cac gcc att gcc atc gac tac gac          1213
Val Leu Gln Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp
                370                     375                     380
ccg cta gag ggc tat gtc tac tgg aca gat gac gag gtg cgg gcc atc          1261
Pro Leu Glu Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile
            385                     390                     395
cgc agg gcg tac ctg gac ggg tct ggg gcg cag acg ctg gtc aac acc          1309
Arg Arg Ala Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr
        400                     405                     410
gag atc aac gac ccc gat ggc atc gcg gtc gac tgg gtg gcc cga aac          1357
Glu Ile Asn Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn
    415                     420                     425
ctc tac tgg acc gac acg ggc acg gac cgc atc gag gtg acg cgc ctc          1405
Leu Tyr Trp Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu
430                 435                     440                     445
aac ggc acc tcc cgc aag atc ctg gtg tcg gag gac ctg gac gag ccc          1453
Asn Gly Thr Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro
                450                     455                     460
cga gcc atc gca ctg cac ccc gtg atg ggc ctc atg tac tgg aca gac          1501
Arg Ala Ile Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp
            465                     470                     475
tgg gga gag aac cct aaa atc gag tgt gcc aac ttg gat ggg cag gag          1549
Trp Gly Glu Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu
        480                     485                     490
cgg cgt gtg ctg gtc aat gcc tcc ctc ggg tgg ccc aac ggc ctg gcc          1597
Arg Arg Val Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala
    495                     500                     505
ctg gac ctg cag gag ggg aag ctc tac tgg gga gac gcc aag aca gac          1645
```

```
Leu Asp Leu Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp
510                 515                 520                 525
aag atc gag gtg atc aat gtt gat ggg acg aag agg cgg acc ctc ctg        1693
Lys Ile Glu Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu
                530                 535                 540
gag gac aag ctc ccg cac att ttc ggg ttc acg ctg ctg ggg gac ttc        1741
Glu Asp Lys Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe
            545                 550                 555
atc tac tgg act gac tgg cag cgc cgc agc atc gag cgg gtg cac aag        1789
Ile Tyr Trp Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys
        560                 565                 570
gtc aag gcc agc cgg gac gtc atc att gac cag ctg ccc gac ctg atg        1837
Val Lys Ala Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met
    575                 580                 585
ggg ctc aaa gct gtg aat gtg gcc aag gtc gtc gga acc aac ccg tgt        1885
Gly Leu Lys Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys
590                 595                 600                 605
gcg gac agg aac ggg ggg tgc agc cac ctg tgc ttc ttc aca ccc cac        1933
Ala Asp Arg Asn Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro His
                610                 615                 620
gca acc cgg tgt ggc tgc ccc atc ggc ctg gag ctg ctg agt gac atg        1981
Ala Thr Arg Cys Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met
            625                 630                 635
aag acc tgc atc gtg cct gag gcc ttc ttg gtc ttc acc agc aga gcc        2029
Lys Thr Cys Ile Val Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala
        640                 645                 650
gcc atc cac agg atc tcc ctc gag acc aat aac aac gac gtg gcc atc        2077
Ala Ile His Arg Ile Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile
    655                 660                 665
ccg ctc acg ggc gtc aag gag gcc tca gcc ctg gac ttt gat gtg tcc        2125
Pro Leu Thr Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser
670                 675                 680                 685
aac aac cac atc tac tgg aca gac gtc agc ctg aag acc atc agc cgc        2173
Asn Asn His Ile Tyr Trp Thr Asp Val Ser Leu Lys Thr Ile Ser Arg
                690                 695                 700
gcc ttc atg aac ggg agc tcg gtg gag cac gtg gtg gag ttt ggc ctt        2221
Ala Phe Met Asn Gly Ser Ser Val Glu His Val Val Glu Phe Gly Leu
            705                 710                 715
gac tac ccc gag ggc atg gcc gtt gac tgg atg ggc aag aac ctc tac        2269
Asp Tyr Pro Glu Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr
        720                 725                 730
tgg gcc gac act ggg acc aac aga atc gaa gtg gcg cgg ctg gac ggg        2317
Trp Ala Asp Thr Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly
    735                 740                 745
cag ttc cgg caa gtc ctc gtg tgg agg gac ttg gac aac ccg agg tcg        2365
Gln Phe Arg Gln Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser
750                 755                 760                 765
ctg gcc ctg gat ccc acc aag ggc tac atc tac tgg acc gag tgg ggc        2413
Leu Ala Leu Asp Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly
                770                 775                 780
ggc aag ccg agg atc gtg cgg gcc ttc atg gac ggg acc aac tgc atg        2461
Gly Lys Pro Arg Ile Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met
            785                 790                 795
acg ctg gtg gac aag gtg ggc cgg gcc aac gac ctc acc att gac tac        2509
Thr Leu Val Asp Lys Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr
        800                 805                 810
gct gac cag cgc ctc tac tgg acc gac ctg gac acc aac atg atc gag        2557
```

```
Ala Asp Gln Arg Leu Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu
    815                 820                 825
tcg tcc aac atg ctg ggt cag gag cgg gtc gtg att gcc gac gat ctc      2605
Ser Ser Asn Met Leu Gly Gln Glu Arg Val Val Ile Ala Asp Asp Leu
830                 835                 840                 845
ccg cac ccg ttc ggt ctg acg cag tac agc gat tat atc tac tgg aca      2653
Pro His Pro Phe Gly Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr
                850                 855                 860
gac tgg aat ctg cac agc att gag cgg gcc gac aag act agc ggc cgg      2701
Asp Trp Asn Leu His Ser Ile Glu Arg Ala Asp Lys Thr Ser Gly Arg
            865                 870                 875
aac cgc acc ctc atc cag ggc cac ctg gac ttc gtg atg gac atc ctg      2749
Asn Arg Thr Leu Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu
        880                 885                 890
gtg ttc cac tcc tcc cgc cag gat ggc ctc aat gac tgt atg cac aac      2797
Val Phe His Ser Ser Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn
    895                 900                 905
aac ggg cag tgt ggg cag ctg tgc ctt gcc atc ccc ggc ggc cac cgc      2845
Asn Gly Gln Cys Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg
910                 915                 920                 925
tgc ggc tgc gcc tca cac tac acc ctg gac ccc agc agc cgc aac tgc      2893
Cys Gly Cys Ala Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys
                930                 935                 940
agc ccg ccc acc acc ttc ttg ctg ttc agc cag aaa tct gcc atc agt      2941
Ser Pro Pro Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser
            945                 950                 955
cgg atg atc ccg gac gac cag cac agc ccg gat ctc atc ctg ccc ctg      2989
Arg Met Ile Pro Asp Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu
        960                 965                 970
cat gga ctg agg aac gtc aaa gcc atc gac tat gac cca ctg gac aag      3037
His Gly Leu Arg Asn Val Lys Ala Ile Asp Tyr Asp Pro Leu Asp Lys
    975                 980                 985
ttc atc tac tgg gtg gat ggg cgc cag aac atc aag cga gcc aag gac      3085
Phe Ile Tyr Trp Val Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp
990                 995                 1000                1005
gac ggg acc cag ccc ttt gtt ttg acc tct ctg agc caa ggc caa aac      3133
Asp Gly Thr Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn
                1010                1015                1020
cca gac agg cag ccc cac gac ctc agc atc gac atc tac agc cgg aca      3181
Pro Asp Arg Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr
            1025                1030                1035
ctg ttc tgg acg tgc gag gcc acc aat acc atc aac gtc cac agg ctg      3229
Leu Phe Trp Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu
        1040                1045                1050
agc ggg gaa gcc atg ggg gtg gtg ctg cgt ggg gac cgc gac aag ccc      3277
Ser Gly Glu Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro
    1055                1060                1065
agg gcc atc gtc gtc aac gcg gag cga ggg tac ctg tac ttc acc aac      3325
Arg Ala Ile Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn
1070                1075                1080                1085
atg cag gac cgg gca gcc aag atc gaa cgc gca gcc ctg gac ggc acc      3373
Met Gln Asp Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr
                1090                1095                1100
gag cgc gag gtc ctc ttc acc acc ggc ctc atc cgc cct gtg gcc ctg      3421
Glu Arg Glu Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu
            1105                1110                1115
gtg gtg gac aac aca ctg ggc aag ctg ttc tgg gtg gac gcg gac ctg      3469
```

```
Val Val Asp Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu
        1120                1125                1130
aag cgc att gag agc tgt gac ctg tca ggg gcc aac cgc ctg acc ctg      3517
Lys Arg Ile Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu
    1135                1140                1145
gag gac gcc aac atc gtg cag cct ctg ggc ctg acc atc ctt ggc aag      3565
Glu Asp Ala Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys
1150                1155                1160                1165
cat ctc tac tgg atc gac cgc cag cag cag atg atc gag cgt gtg gag      3613
His Leu Tyr Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu
                1170                1175                1180
aag acc acc ggg gac aag cgg act cgc atc cag ggc cgt gtc gcc cac      3661
Lys Thr Thr Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His
            1185                1190                1195
ctc act ggc atc cat gca gtg gag gaa gtc agc ctg gag gag ttc tca      3709
Leu Thr Gly Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser
        1200                1205                1210
gcc cac cca tgt gcc cgt gac aat ggt ggc tgc tcc cac atc tgt att      3757
Ala His Pro Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile
    1215                1220                1225
gcc aag ggt gat ggg aca cca cgg tgc tca tgc cca gtc cac ctc gtg      3805
Ala Lys Gly Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val
1230                1235                1240                1245
ctc ctg cag aac ctg ctg acc tgt gga gag ccg ccc acc tgc tcc ccg      3853
Leu Leu Gln Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro
                1250                1255                1260
gac cag ttt gca tgt gcc aca ggg gag atc gac tgt atc ccc ggg gcc      3901
Asp Gln Phe Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala
            1265                1270                1275
tgg cgc tgt gac ggc ttt ccc gag tgc gat gac cag agc gac gag gag      3949
Trp Arg Cys Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu
        1280                1285                1290
ggc tgc ccc gtg tgc tcc gcc gcc cag ttc ccc tgc gcg cgg ggt cag      3997
Gly Cys Pro Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln
    1295                1300                1305
tgt gtg gac ctg cgc ctg cgc tgc gac ggc gag gca gac tgt cag gac      4045
Cys Val Asp Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp
1310                1315                1320                1325
cgc tca gac gag gtg gac tgt gac gcc atc tgc ctg ccc aac cag ttc      4093
Arg Ser Asp Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe
                1330                1335                1340
cgg tgt gcg agc ggc cag tgt gtc ctc atc aaa cag cag tgc gac tcc      4141
Arg Cys Ala Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser
            1345                1350                1355
ttc ccc gac tgt atc gac ggc tcc gac gag ctc atg tgt gaa atc acc      4189
Phe Pro Asp Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr
        1360                1365                1370
aag ccg ccc tca gac gac agc ccg gcc cac agc agt gcc atc ggg ccc      4237
Lys Pro Pro Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro
    1375                1380                1385
gtc att ggc atc atc ctc tct ctc ttc gtc atg ggt ggt gtc tat ttt      4285
Val Ile Gly Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe
1390                1395                1400                1405
gtg tgc cag cgc gtg gtg tgc cag cgc tat gcg ggg gcc aac ggg ccc      4333
Val Cys Gln Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro
                1410                1415                1420
ttc ccg cac gag tat gtc agc ggg acc ccg cac gtg ccc ctc aat ttc      4381
```

```
 Phe Pro His Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe
             1425                1430                1435
 ata gcc ccg ggc ggt tcc cag cat ggc ccc ttc aca ggc atc gca tgc        4429
 Ile Ala Pro Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys
         1440                1445                1450
 gga aag tcc atg atg agc tcc gtg agc ctg atg ggg ggc cgg ggc ggg        4477
 Gly Lys Ser Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly
     1455                1460                1465
 gtg ccc ctc tac gac cgg aac cac gtc aca ggg gcc tcg tcc agc agc        4525
 Val Pro Leu Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser
 1470                1475                1480                1485
 tcg tcc agc acg aag gcc acg ctg tac ccg ccg atc ctg aac ccg ccg        4573
 Ser Ser Ser Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro
                 1490                1495                1500
 ccc tcc ccg gcc acg gac ccc tcc ctg tac aac atg gac atg ttc tac        4621
 Pro Ser Pro Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr
             1505                1510                1515
 tct tca aac att ccg gcc act gcg aga ccg tac agg ccc tac atc att        4669
 Ser Ser Asn Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile
         1520                1525                1530
 cga gga atg gcg ccc ccg acg acg ccc tgc agc acc gac gtg tgt gac        4717
 Arg Gly Met Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp
     1535                1540                1545
 agc gac tac agc gcc agc cgc tgg aag gcc agc aag tac tac ctg gat        4765
 Ser Asp Tyr Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp
 1550                1555                1560                1565
 ttg aac tcg gac tca gac ccc tat cca ccc cca ccc acg ccc cac agc        4813
 Leu Asn Ser Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser
                 1570                1575                1580
 cag tac ctg tcg gcg gag gac agc tgc ccg ccc tcg ccc gcc acc gag        4861
 Gln Tyr Leu Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu
             1585                1590                1595
 agg agc tac ttc cat ctc ttc ccg ccc cct ccg tcc ccc tgc acg gac        4909
 Arg Ser Tyr Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp
         1600                1605                1610
 tca tcc tgacctcggc cgggccactc tggcttctct gtgcccctgt aaatagtttt         4965
 Ser Ser
     1615
aaatatgaac aaagaaaaaa atatatttta tgatttaaaa aataaatata attgggattt       5025
taaaaacatg agaaatgtga actgtgatgg ggtgggcagg gctgggagaa ctttgtacag       5085
tggagaaata tttataaact taattttgta aaaca                                  5120


&lt;210&gt; 3
&lt;211&gt; 1615
&lt;212&gt; PRT
&lt;213&gt; Homo sapiens


&lt;400&gt; 3
Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu Leu Leu Leu
1               5                   10                  15
Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala Ala Ala Ser
            20                  25                  30
Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu Val Asp Ala
        35                  40                  45
Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly Leu Glu Asp
```

50 55 60
Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val Tyr Trp Thr
65 70 75 80
Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn Gln Thr Gly
85 90 95
Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser Pro Asp Gly
100 105 110
Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr Asp Ser Glu
115 120 125
Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser Arg Lys Val
130 135 140
Leu Phe Trp Gln Asp Leu Asp Gln Pro Lys Ala Ile Ala Leu Asp Pro
145 150 155 160
Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Gly Glu Thr Pro Arg Ile
165 170 175
Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile Val Asp Ser
180 185 190
Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu Glu Gln Lys
195 200 205
Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg Ala Asn Leu
210 215 220
Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu Thr His Pro
225 230 235 240
Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr Asp Trp Gln
245 250 255
Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly Lys Arg Lys
260 265 270
Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln Val Leu Ser
275 280 285
Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu Asp Asn Gly
290 295 300
Gly Trp Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro Phe Tyr Thr
305 310 315 320
Cys Ala Cys Pro Thr Gly Val Gln Met Gln Asp Asn Gly Arg Thr Cys
325 330 335
Lys Ala Gly Ala Glu Glu Val Leu Leu Leu Ala Arg Arg Thr Asp Leu
340 345 350
Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile Val Leu Gln
355 360 365
Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp Pro Leu Glu
370 375 380
Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile Arg Arg Ala
385 390 395 400
Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr Glu Ile Asn
405 410 415
Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn Leu Tyr Trp
420 425 430
Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu Asn Gly Thr
435 440 445
Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro Arg Ala Ile
450 455 460
Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp Trp Gly Glu
465 470 475 480
Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu Arg Arg Val
485 490 495
Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala Leu Asp Leu
500 505 510

```
Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp Lys Ile Glu
        515                 520                 525
Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu Glu Asp Lys
    530                 535                 540
Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe Ile Tyr Trp
545                 550                 555                 560
Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys Val Lys Ala
                565                 570                 575
Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met Gly Leu Lys
            580                 585                 590
Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys Ala Asp Arg
        595                 600                 605
Asn Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro His Ala Thr Arg
    610                 615                 620
Cys Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met Lys Thr Cys
625                 630                 635                 640
Ile Val Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala Ala Ile His
                645                 650                 655
Arg Ile Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile Pro Leu Thr
            660                 665                 670
Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser Asn Asn His
        675                 680                 685
Ile Tyr Trp Thr Asp Val Ser Leu Lys Asn Ile Ser Arg Ala Phe Met
    690                 695                 700
Asn Gly Ser Ser Val Glu His Val Val Glu Phe Gly Leu Asp Tyr Pro
705                 710                 715                 720
Glu Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr Trp Ala Asp
                725                 730                 735
Thr Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly Gln Phe Arg
            740                 745                 750
Gln Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser Leu Ala Leu
        755                 760                 765
Asp Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly Gly Lys Pro
    770                 775                 780
Arg Ile Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met Thr Leu Val
785                 790                 795                 800
Asp Lys Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr Ala Asp Gln
                805                 810                 815
Arg Leu Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu Ser Ser Asn
            820                 825                 830
Met Leu Gly Gln Glu Arg Val Val Ile Ala Asp Asp Leu Pro His Pro
        835                 840                 845
Phe Gly Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr Asp Trp Asn
    850                 855                 860
Leu His Ser Ile Glu Arg Ala Asp Lys Thr Ser Gly Arg Asn Arg Thr
865                 870                 875                 880
Leu Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu Val Phe His
                885                 890                 895
Ser Ser Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn Asn Gly Gln
            900                 905                 910
Cys Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg Cys Gly Cys
        915                 920                 925
Ala Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys Ser Pro Pro
    930                 935                 940
Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser Arg Met Ile
945                 950                 955                 960
Pro Asp Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu His Gly Leu
```

```
                965                 970                 975
Arg Asn Val Lys Ala Ile Asp Tyr Asp Pro Leu Asp Lys Phe Ile Tyr
            980                 985                 990
Trp Val Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp Asp Gly Thr
        995                 1000                1005
Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn Pro Asp Arg
    1010                1015                1020
Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr Leu Phe Trp
1025                1030                1035                1040
Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu Ser Gly Glu
                1045                1050                1055
Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro Arg Ala Ile
            1060                1065                1070
Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn Met Gln Asp
        1075                1080                1085
Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr Glu Arg Glu
    1090                1095                1100
Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu Val Val Asp
1105                1110                1115                1120
Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu Lys Arg Ile
                1125                1130                1135
Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu Glu Asp Ala
            1140                1145                1150
Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys His Leu Tyr
        1155                1160                1165
Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu Lys Thr Thr
    1170                1175                1180
Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His Leu Thr Gly
1185                1190                1195                1200
Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser Ala His Pro
                1205                1210                1215
Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile Ala Lys Gly
            1220                1225                1230
Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val Leu Leu Gln
        1235                1240                1245
Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro Asp Gln Phe
    1250                1255                1260
Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala Trp Arg Cys
1265                1270                1275                1280
Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu Gly Cys Pro
                1285                1290                1295
Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln Cys Val Asp
            1300                1305                1310
Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp Arg Ser Asp
        1315                1320                1325
Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe Arg Cys Ala
    1330                1335                1340
Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser Phe Pro Asp
1345                1350                1355                1360
Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr Lys Pro Pro
                1365                1370                1375
Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro Val Ile Gly
            1380                1385                1390
Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe Val Cys Gln
        1395                1400                1405
Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro Phe Pro His
    1410                1415                1420
```

Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe Ile Ala Pro
1425 1430 1435 1440
Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys Gly Lys Ser
1445 1450 1455
Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly Val Pro Leu
1460 1465 1470
Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser Ser Ser Ser
1475 1480 1485
Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro Pro Ser Pro
1490 1495 1500
Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr Ser Ser Asn
1505 1510 1515 1520
Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile Arg Gly Met
1525 1530 1535
Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp Ser Asp Tyr
1540 1545 1550
Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp Leu Asn Ser
1555 1560 1565
Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser Gln Tyr Leu
1570 1575 1580
Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu Arg Ser Tyr
1585 1590 1595 1600
Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp Ser Ser
1605 1610 1615

<210> 4
<211> 1615
<212> PRT
<213> Homo sapiens

<400> 4
Met Glu Ala Ala Pro Pro Gly Pro Pro Trp Pro Leu Leu Leu Leu Leu
1 5 10 15
Leu Leu Leu Leu Ala Leu Cys Gly Cys Pro Ala Pro Ala Ala Ala Ser
20 25 30
Pro Leu Leu Leu Phe Ala Asn Arg Arg Asp Val Arg Leu Val Asp Ala
35 40 45
Gly Gly Val Lys Leu Glu Ser Thr Ile Val Val Ser Gly Leu Glu Asp
50 55 60
Ala Ala Ala Val Asp Phe Gln Phe Ser Lys Gly Ala Val Tyr Trp Thr
65 70 75 80
Asp Val Ser Glu Glu Ala Ile Lys Gln Thr Tyr Leu Asn Gln Thr Gly
85 90 95
Ala Ala Val Gln Asn Val Val Ile Ser Gly Leu Val Ser Pro Asp Gly
100 105 110
Leu Ala Cys Asp Trp Val Gly Lys Lys Leu Tyr Trp Thr Asp Ser Glu
115 120 125
Thr Asn Arg Ile Glu Val Ala Asn Leu Asn Gly Thr Ser Arg Lys Val
130 135 140
Leu Phe Trp Gln Asp Leu Asp Gln Pro Lys Ala Ile Ala Leu Asp Pro
145 150 155 160
Ala His Gly Tyr Met Tyr Trp Thr Asp Trp Val Glu Thr Pro Arg Ile
165 170 175
Glu Arg Ala Gly Met Asp Gly Ser Thr Arg Lys Ile Ile Val Asp Ser
180 185 190

Asp Ile Tyr Trp Pro Asn Gly Leu Thr Ile Asp Leu Glu Glu Gln Lys
        195   200   205
Leu Tyr Trp Ala Asp Ala Lys Leu Ser Phe Ile His Arg Ala Asn Leu
    210   215   220
Asp Gly Ser Phe Arg Gln Lys Val Val Glu Gly Ser Leu Thr His Pro
225   230   235   240
Phe Ala Leu Thr Leu Ser Gly Asp Thr Leu Tyr Trp Thr Asp Trp Gln
                245   250   255
Thr Arg Ser Ile His Ala Cys Asn Lys Arg Thr Gly Gly Lys Arg Lys
            260   265   270
Glu Ile Leu Ser Ala Leu Tyr Ser Pro Met Asp Ile Gln Val Leu Ser
        275   280   285
Gln Glu Arg Gln Pro Phe Phe His Thr Arg Cys Glu Glu Asp Asn Gly
    290   295   300
Gly Trp Ser His Leu Cys Leu Leu Ser Pro Ser Glu Pro Phe Tyr Thr
305   310   315   320
Cys Ala Cys Pro Thr Gly Val Gln Met Gln Asp Asn Gly Arg Thr Cys
                325   330   335
Lys Ala Gly Ala Glu Glu Val Leu Leu Leu Ala Arg Arg Thr Asp Leu
            340   345   350
Arg Arg Ile Ser Leu Asp Thr Pro Asp Phe Thr Asp Ile Val Leu Gln
        355   360   365
Val Asp Asp Ile Arg His Ala Ile Ala Ile Asp Tyr Asp Pro Leu Glu
    370   375   380
Gly Tyr Val Tyr Trp Thr Asp Asp Glu Val Arg Ala Ile Arg Arg Ala
385   390   395   400
Tyr Leu Asp Gly Ser Gly Ala Gln Thr Leu Val Asn Thr Glu Ile Asn
                405   410   415
Asp Pro Asp Gly Ile Ala Val Asp Trp Val Ala Arg Asn Leu Tyr Trp
            420   425   430
Thr Asp Thr Gly Thr Asp Arg Ile Glu Val Thr Arg Leu Asn Gly Thr
        435   440   445
Ser Arg Lys Ile Leu Val Ser Glu Asp Leu Asp Glu Pro Arg Ala Ile
    450   455   460
Ala Leu His Pro Val Met Gly Leu Met Tyr Trp Thr Asp Trp Gly Glu
465   470   475   480
Asn Pro Lys Ile Glu Cys Ala Asn Leu Asp Gly Gln Glu Arg Arg Val
                485   490   495
Leu Val Asn Ala Ser Leu Gly Trp Pro Asn Gly Leu Ala Leu Asp Leu
            500   505   510
Gln Glu Gly Lys Leu Tyr Trp Gly Asp Ala Lys Thr Asp Lys Ile Glu
        515   520   525
Val Ile Asn Val Asp Gly Thr Lys Arg Arg Thr Leu Leu Glu Asp Lys
    530   535   540
Leu Pro His Ile Phe Gly Phe Thr Leu Leu Gly Asp Phe Ile Tyr Trp
545   550   555   560
Thr Asp Trp Gln Arg Arg Ser Ile Glu Arg Val His Lys Val Lys Ala
                565   570   575
Ser Arg Asp Val Ile Ile Asp Gln Leu Pro Asp Leu Met Gly Leu Lys
            580   585   590
Ala Val Asn Val Ala Lys Val Val Gly Thr Asn Pro Cys Ala Asp Arg
        595   600   605
Asn Gly Gly Cys Ser His Leu Cys Phe Phe Thr Pro His Ala Thr Arg
    610   615   620
Cys Gly Cys Pro Ile Gly Leu Glu Leu Leu Ser Asp Met Lys Thr Cys
625   630   635   640
Ile Val Pro Glu Ala Phe Leu Val Phe Thr Ser Arg Ala Ala Ile His

```
                645                 650                 655
Arg Ile Ser Leu Glu Thr Asn Asn Asn Asp Val Ala Ile Pro Leu Thr
            660                 665                 670
Gly Val Lys Glu Ala Ser Ala Leu Asp Phe Asp Val Ser Asn Asn His
        675                 680                 685
Ile Tyr Trp Thr Asp Val Ser Leu Lys Asn Ile Ser Arg Ala Phe Met
    690                 695                 700
Asn Gly Ser Ser Val Glu His Val Val Glu Phe Gly Leu Asp Tyr Pro
705                 710                 715                 720
Glu Gly Met Ala Val Asp Trp Met Gly Lys Asn Leu Tyr Trp Ala Asp
                725                 730                 735
Thr Gly Thr Asn Arg Ile Glu Val Ala Arg Leu Asp Gly Gln Phe Arg
            740                 745                 750
Gln Val Leu Val Trp Arg Asp Leu Asp Asn Pro Arg Ser Leu Ala Leu
        755                 760                 765
Asp Pro Thr Lys Gly Tyr Ile Tyr Trp Thr Glu Trp Gly Gly Lys Pro
    770                 775                 780
Arg Ile Val Arg Ala Phe Met Asp Gly Thr Asn Cys Met Thr Leu Val
785                 790                 795                 800
Asp Lys Val Gly Arg Ala Asn Asp Leu Thr Ile Asp Tyr Ala Asp Gln
                805                 810                 815
Arg Leu Tyr Trp Thr Asp Leu Asp Thr Asn Met Ile Glu Ser Ser Asn
            820                 825                 830
Met Leu Gly Gln Glu Arg Val Val Ile Ala Asp Asp Leu Pro His Pro
        835                 840                 845
Phe Gly Leu Thr Gln Tyr Ser Asp Tyr Ile Tyr Trp Thr Asp Trp Asn
    850                 855                 860
Leu His Ser Ile Glu Arg Ala Asp Lys Thr Ser Gly Arg Asn Arg Thr
865                 870                 875                 880
Leu Ile Gln Gly His Leu Asp Phe Val Met Asp Ile Leu Val Phe His
                885                 890                 895
Ser Ser Arg Gln Asp Gly Leu Asn Asp Cys Met His Asn Asn Gly Gln
            900                 905                 910
Cys Gly Gln Leu Cys Leu Ala Ile Pro Gly Gly His Arg Cys Gly Cys
        915                 920                 925
Ala Ser His Tyr Thr Leu Asp Pro Ser Ser Arg Asn Cys Ser Pro Pro
    930                 935                 940
Thr Thr Phe Leu Leu Phe Ser Gln Lys Ser Ala Ile Ser Arg Met Ile
945                 950                 955                 960
Pro Asp Asp Gln His Ser Pro Asp Leu Ile Leu Pro Leu His Gly Leu
                965                 970                 975
Arg Asn Val Lys Ala Ile Asp Tyr Asp Pro Leu Asp Lys Phe Ile Tyr
            980                 985                 990
Trp Val Asp Gly Arg Gln Asn Ile Lys Arg Ala Lys Asp Asp Gly Thr
        995                 1000                1005
Gln Pro Phe Val Leu Thr Ser Leu Ser Gln Gly Gln Asn Pro Asp Arg
    1010                1015                1020
Gln Pro His Asp Leu Ser Ile Asp Ile Tyr Ser Arg Thr Leu Phe Trp
1025                1030                1035                1040
Thr Cys Glu Ala Thr Asn Thr Ile Asn Val His Arg Leu Ser Gly Glu
                1045                1050                1055
Ala Met Gly Val Val Leu Arg Gly Asp Arg Asp Lys Pro Arg Ala Ile
            1060                1065                1070
Val Val Asn Ala Glu Arg Gly Tyr Leu Tyr Phe Thr Asn Met Gln Asp
        1075                1080                1085
Arg Ala Ala Lys Ile Glu Arg Ala Ala Leu Asp Gly Thr Glu Arg Glu
    1090                1095                1100
```

Val Leu Phe Thr Thr Gly Leu Ile Arg Pro Val Ala Leu Val Val Asp
1105 1110 1115 1120
Asn Thr Leu Gly Lys Leu Phe Trp Val Asp Ala Asp Leu Lys Arg Ile
1125 1130 1135
Glu Ser Cys Asp Leu Ser Gly Ala Asn Arg Leu Thr Leu Glu Asp Ala
1140 1145 1150
Asn Ile Val Gln Pro Leu Gly Leu Thr Ile Leu Gly Lys His Leu Tyr
1155 1160 1165
Trp Ile Asp Arg Gln Gln Gln Met Ile Glu Arg Val Glu Lys Thr Thr
1170 1175 1180
Gly Asp Lys Arg Thr Arg Ile Gln Gly Arg Val Ala His Leu Thr Gly
1185 1190 1195 1200
Ile His Ala Val Glu Glu Val Ser Leu Glu Glu Phe Ser Ala His Pro
1205 1210 1215
Cys Ala Arg Asp Asn Gly Gly Cys Ser His Ile Cys Ile Ala Lys Gly
1220 1225 1230
Asp Gly Thr Pro Arg Cys Ser Cys Pro Val His Leu Val Leu Leu Gln
1235 1240 1245
Asn Leu Leu Thr Cys Gly Glu Pro Pro Thr Cys Ser Pro Asp Gln Phe
1250 1255 1260
Ala Cys Ala Thr Gly Glu Ile Asp Cys Ile Pro Gly Ala Trp Arg Cys
1265 1270 1275 1280
Asp Gly Phe Pro Glu Cys Asp Asp Gln Ser Asp Glu Glu Gly Cys Pro
1285 1290 1295
Val Cys Ser Ala Ala Gln Phe Pro Cys Ala Arg Gly Gln Cys Val Asp
1300 1305 1310
Leu Arg Leu Arg Cys Asp Gly Glu Ala Asp Cys Gln Asp Arg Ser Asp
1315 1320 1325
Glu Val Asp Cys Asp Ala Ile Cys Leu Pro Asn Gln Phe Arg Cys Ala
1330 1335 1340
Ser Gly Gln Cys Val Leu Ile Lys Gln Gln Cys Asp Ser Phe Pro Asp
1345 1350 1355 1360
Cys Ile Asp Gly Ser Asp Glu Leu Met Cys Glu Ile Thr Lys Pro Pro
1365 1370 1375
Ser Asp Asp Ser Pro Ala His Ser Ser Ala Ile Gly Pro Val Ile Gly
1380 1385 1390
Ile Ile Leu Ser Leu Phe Val Met Gly Gly Val Tyr Phe Val Cys Gln
1395 1400 1405
Arg Val Val Cys Gln Arg Tyr Ala Gly Ala Asn Gly Pro Phe Pro His
1410 1415 1420
Glu Tyr Val Ser Gly Thr Pro His Val Pro Leu Asn Phe Ile Ala Pro
1425 1430 1435 1440
Gly Gly Ser Gln His Gly Pro Phe Thr Gly Ile Ala Cys Gly Lys Ser
1445 1450 1455
Met Met Ser Ser Val Ser Leu Met Gly Gly Arg Gly Gly Val Pro Leu
1460 1465 1470
Tyr Asp Arg Asn His Val Thr Gly Ala Ser Ser Ser Ser Ser Ser Ser
1475 1480 1485
Thr Lys Ala Thr Leu Tyr Pro Pro Ile Leu Asn Pro Pro Pro Ser Pro
1490 1495 1500
Ala Thr Asp Pro Ser Leu Tyr Asn Met Asp Met Phe Tyr Ser Ser Asn
1505 1510 1515 1520
Ile Pro Ala Thr Ala Arg Pro Tyr Arg Pro Tyr Ile Ile Arg Gly Met
1525 1530 1535
Ala Pro Pro Thr Thr Pro Cys Ser Thr Asp Val Cys Asp Ser Asp Tyr
1540 1545 1550
Ser Ala Ser Arg Trp Lys Ala Ser Lys Tyr Tyr Leu Asp Leu Asn Ser

```
        1555                1560                1565
Asp Ser Asp Pro Tyr Pro Pro Pro Pro Thr Pro His Ser Gln Tyr Leu
    1570                1575                1580
Ser Ala Glu Asp Ser Cys Pro Pro Ser Pro Ala Thr Glu Arg Ser Tyr
1585                1590                1595                1600
Phe His Leu Phe Pro Pro Pro Pro Ser Pro Cys Thr Asp Ser Ser
                1605                1610                1615
```

<210> 5
<211> 3096
<212> DNA
<213> Homo sapiens

<400> 5

```
catcttctca cacgatctct cgcttcgcac tccttccttt gattggtttt caccatttac     60
tcagacgacg gtccttcttc gatctttgca cattcttcta tcatctacta ccttcatacc    120
cagctccgtc ccctaatatt catgcgcgga tggcccattc cgtggtgaaa attcccttct    180
actctgctaa tctgctgttc tctctccctc ccgtcgggtt ctgctcctgc cacgttctcc    240
cctctcccca ccaaaggctg ggttttcttt gtcagggctc ctttcccctt tggaagaagg    300
ggggctgtat ggccttggtg cgaggccctc cagtgacagg atcccccatc acccagagtt    360
ccacaggccc tggtagggag gagggggagc agaagaggag gtgccatctt tgcctgctgg    420
ggaagggcag gggccaccca cacagagctc tcccatttgc tgtggaccct ggggccactg    480
cccagttcct tccaaaggaa agccagctcc ccaggtggtg ggagagtgat atggcttcct    540
cttaaactta gggaattgag tgtgtggttg cttctaagtg ccttagaagc cgggagcggc    600
tcctggaaag agcctgcctg ccacagcggg ccttaccctg gctgtgccca cagatgtccc    660
tggggcctgc cgctcctgcc cggctctcct ggcctccccc ggtgtgggtt gggaaaagca    720
cagcaaatta aaaaacacct ccatctctgg cctttgaaga atgcatctga acagccgaga    780
gtgtaaaccg tggtgaaatg tggtctttcc agtttgggga gaagcagggc agagctgggg    840
cttttgtacc cagggtttcc aagagctcct gcctccctcg gctgggctgg ccagggcccc    900
ccgctgggac ctccagctgt aatagggaag gttttactgg gttgctggcc actgtggact    960
gcccctaagg gcaggtatgc ctgcctttac ccgggttccc ctcctgcctg gaagatacag   1020
cccatgggag gcctgttgtc tgtgggatcc tccagcatca gagacactgg ggccagcgtc   1080
tgcctggtga ggtgcaggcc tggcaggccc ggtcccccac ctgcttgagc acccacggtg   1140
gtgggggctc gctgcctccc gagacaatct atgtcattgt tgtccaagga agctaattta   1200
gagtagaaag ttccgtgtcc agtcccactc tgtgcgtgtg ttagcagggg actctcgggc   1260
cggagctggg tccaccctgg tagggggact tcatggggcc tgggcgacag cactgtgtat   1320
ttgtgtgtgt gtgtgtttgt gtgtgtgtgt gtctgaggag gtggaccagt ttctcaaaag   1380
gcctgtgacc ccaagaacca aggaatttca gcctgggtgg atcacacctt cactggtgag   1440
tgggacaagc tgggggccct cgccacagga gcagccaggg catggggcac agttggcctc   1500
attcacaaaa tgggagtata agtgatccct gctctggcgg ccaggacgat gagtgggaac   1560
acaccgtgtg gggggctgcct ggcctgggtg tgccgcgggt gtccttgttg gtgatggttc   1620
cacctgcttg tgccaccagt gccctctggg tctcacacac aactctcttc ccagcgaagg   1680
cccctcctgc cctcaggcct cagtgctgct tccgtctcgg aaggccccag gagctcctgc   1740
atcctgggcg tgattcctgt gtgcctgcag accccctcgc ggctgccatc tcatcctttg   1800
gtgcacctgt tggccagacc tcctggtagc gggtgctgca ctcccctgaa tgtgccgggg   1860
cctgggggca gggacctggg ctcctccctc actgagtgga gggaactcag tgtcttggag   1920
ttggggtgcc tgcaggctgg gtggtgcagg tgaaatgcag acctctcagc tggtgttcca   1980
gagcagctgc cttcccccgc ccgagggact tcacccgcag cccagtcagg ggtggcgcct   2040
gggtgcatcg cccgcaggct gggtaggggt ggagcctggg tggccctgcc tgtgagctgc   2100
atagttgtcg cctttgaccc tgagttttct tcgttatctg tttggacctg tttggggcag   2160
gcaggggatg agatctgaag ataaatgcct tagctgtgac catctccttt tgtgagaggt   2220
caatgtccag ttccgctgca gttataacat cccatttttt gatttctttt tattttttcc   2280
tttttctttt tgagatggag tctcgctctg tcacccaggc tggagtgcaa tggggtgacc   2340
tcagctcact gcaacctcca cttctcgggt tcaagtgatt ctcctgcctc agcctcctga   2400
```

```
ctagcagggg ttacaggcgt gagccaccac gcccagctaa tttttgtatt tttagtagag    2460
gcaaggtttc gtcatgttgg ccaggctggt ctcaaactcc tggccttaag tgatctgccc    2520
gcctcggcct cccaaagtgc tgagatgaca ggtgtgagcc accgtgcccg gcccagaact    2580
ctttaattcc cacctgaaac ttgccgcctt aagcaggtcc ccagtctccc tcccctagtc    2640
cctggtccca ccattctgct ttctgtctca atgaatttgc ctaccgtaag tacctcatat    2700
aaattgaatc ataaagtatt tgtcttttta tatctggctt atttcactta gcataacatt    2760
cttaagtttc atccatgttg tagcatgtgt cagaatctct ctcttttttt tttttttttt    2820
tttttttttt ttttgcagac agagtctcgc tctgtcatct agactggagt tcagtggcac    2880
gatctcggtt cactgcaaca tctgcctcct gggtccaagc aattctcctg cctcagcctc    2940
cttagcagct ggaactacag gcgcgtgcca ccatgccttg ctaatttttg tatttttatg    3000
tggaggcagg gtttcaccat cttggccagg ctggtctcga attcctggtc ttcaccacgg    3060
gggcccgaag gacccgggca aagcgtggag gggagg                              3096
```

<210> 6
<211> 26928
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> (12044),(12489),(26433),(26434),(26435),(26436),(26439),(26441)
<223> Identity of nucleotide sequences at the above locations are unknown.

<400> 6

```
gaagaccaag ggcacacagc gaggcagttt cagggcgggc agcctggggc cccacggggc    60
ggccccggac acttgttctc acctgtggag ggcagagaag ggaacaggga gagaagtggc    120
cggctgggag tggaggtggg tttgaggttt tactgtaaac taaatgtgta ccctctacct    180
tagttatgaa ttatgagaca cgaagactgc gaaacagaca cactcctcta aaagtgcctc    240
taggctgaca gggagaaagt cccgccaggc tcccagacgc cacctttgag tccttcaaca    300
agcccgccag ggcctcttgc ccaccggtgt cagctcagcc actgaaccct ccaggaagaa    360
gacgtgctgg taggagaaga atctcaccca ggcacagcct ggaaggggca cagaaggggc    420
tccggaacca gcaagcccaa gttggaactc ccagtctgct actttctaga acgactgtgc    480
ccttggcggg tctaagtaga acctctccgc gcactctttc ctcctttgta aagtggggac    540
agcaatggcc accttgcagg ttcagagagg gcttgcagta cctcacagaa ctgagtgccc    600
gtgaacgtgt gtgttcctcc agatttgtga cagctttgcc aggctggagt caggctgaac    660
gcctctgccc tcatggggtt tatattctag gaagaccaac aaaaacaaga agacggaaaa    720
ttaaaacaac aaaagcccca ttgacaggcc gtgaagaatg ccatgaaaaa tgaatggcgt    780
tgtgctgcag tctttgggga aacgggctta cggaaagaag gacacttgag ctgctaccaa    840
tgagcagccg tccggtggga gggcagttca ggaagagcag acatccactg aggaggcgct    900
ggggcagagg gcagcctggt cgctggattc gggggaggaa ccacatcagg ccatgagctg    960
gagctggtgg tagaatgtac aggagaggcc agccagggcc agctcatgtc agacctcaag    1020
cggggaagat gaatcgagaa tgcaccccac gagcaatggg aagccagtct acgatttaag    1080
cagcaaaaat attttccctt cttccaccct gcatccagct ctaccagcac agcctggggt    1140
tctattttca agatagaata gacccagact cccagctctt cttacacttc tactactgcc    1200
acctgtcacc cactcatgcg tccccacttg cagcctcgac ccccttccac ctgatctcat    1260
ggcagccagg gaagctccag ggctcgtgag ggctgccatc tcaggaaaga agcaaaagcc    1320
ttcggcacct gcagggcctg ctccaaccac acttcttcct tgacctctca gcttccttag    1380
ccactccctt cccacatctc accctgctcc agccacagtg gtgtctctgt gggttctcaa    1440
acacaccagg tgcactcctg cctcagggcc tttgtgcttg ctgttctctg ctgggactct    1500
tttttttttt ttttttttg agacagggtc tcactctgtg gcccaggctg gagtgtagtg    1560
gtgtgatcgt agctcattgc aacctcaaac tcctgggctc aagcaatcct cccacctcag    1620
cctctcaagt agttagcttt tgttgttttg ttttgagatg ggatctcact ctgttgccca    1680
ggctggagtg cagtggggca atcttggctc accacaacct ctgcctccca ggctcaagca    1740
attctcctgc ctcagcctcc caagtagctg ggattacagg catgtgccac cacgcccagc    1800
```

```
ttatttttgt attttagta gagacagggt ttcaccatgt tggtctggct ggtcttgaac     1860
tcctggcctc agatgatcca cctgcctcgg cctcccaaag tgctgggatg acaggcatga     1920
gcctgtctct agtagttagg actacagaga ggggccatca tgcctggtga tcctcccacc     1980
ttttctgctc caactctttc accccactta gcctcgtggc tcactctctt acctcttcag     2040
ctcctcagtc aggcctgagg acccctgttg aaaattgcaa accacacccc ccaccaccac     2100
cacccactat tgccagcact ttctactcca tttctctgct ttacttttct cctttgtact     2160
catcaccacc tgactcatta catgtttacg tatctttctt ctctccacta gcatggaagc     2220
tccaggagag cagagagtgt agttttattc cctgatgtgt ttcctgtgcc cgtaccaggg     2280
cctagcacac agtaggtgct cagtaaatgt gtgttggatg aacaaataca gtgaaaggat     2340
ctgatctaca tttataaaga aggcactctg gctgctgagt ggggatgaga ctgtcaggag     2400
gaaagaggcc cctgtggggg cctggccagc aggtgggtac aatggtagca gccaggagag     2460
agggcctctt ggactcaagt ggatggggcc tgctcagggc tccggccaca ggaacaaagg     2520
gaagggggcc caggatggcc tgtcatagag gacacattac aactggccca aagttcaagt     2580
caggtttcta aatttgggaa gggatacaga aaaactaaag actctactgg acagtcagtt     2640
attgaaatga ttacatagaa aatgtaccaa gaattaaaaa aaaaaaaaaa aagcattatg     2700
aaggggccac cagagactcc cagagaggaa agggactatg ggctggatgc ggtgactcac     2760
acctataatc ccagcacttt gggaggccga ggagggtgga tcacgaggtc aggagttcaa     2820
aaccagccta ggcaacatgg taaaaccccc gtttctacta aaaatacaaa aaattagctg     2880
ggcatggcag catgtgcctg taatcccagc tactcgggag gctgaggcag gagagttgct     2940
agaacccagg aggcagaggt tgcagtgagc cgagattgag ccactatgct ccagcttggg     3000
cgacagagca agactccgtc tctaaaaaaa agaaaaaaaa ggccagatga ggtggctcat     3060
gcctgtaatc ccagcacttt gggaggccga ggtgggtgga tcacgaggtc aggagatcga     3120
gaccatcctg gctaacatgg tgaaactcca tctctactta aaatacaaaa aattagccgg     3180
gcgtggtggc gggcacctgt agtcccagct acttgggagg ctgaggcagg agaatggcgt     3240
gaacctggga ggcggagctt gcagtgagcc gagattgcgc cactgcactc catccagcct     3300
gggcgacaga gttagactcc gtctcaaaaa aaaaaaaaaa aaaaaaatta gctgattagt     3360
tgggcttggt ggcgggcgcc tgtaatccca actactcggg aggctgaggc gggagaatca     3420
cttgaacccg ggaggcagag gttgcaatga gccgatatca cgccactaca ctccagcctg     3480
ggcgacagag caagactcca tctcaaaaaa gaaaaaaaaa aagaaagggg ctgtgctgtg     3540
gcctgggacc caaagcacac tactgcaagg tcccagggtg cctgactcca accggagcct     3600
tgagaacatt catttgcaaa gaatgaatta aaattcagca ctattttatt ctgcaggatt     3660
ccagcacccc aaggacagtc atttttagac ccttcagtaa cgtaataagt aaccggagga     3720
tgtgctgagc ttccacttcc ccagacggtt gcctgtcaca gctcatcagg ccaacaaact     3780
tttcttaggc ctcaaatttg gaaatgttca ctctcagttc gttccttaga tgcaagtcca     3840
tcccaatgaa gtaacagggg ctcagcacct gtccaatctc attgcttccg gggacagggg     3900
cccatgagga tgtcgtttca gcccggtgac acttgggcaa agtgcctttt ggtttccctc     3960
ccaggctgga acgtgctggc tctgtgaagt tacgctgggc acaagagccc cccccaaccc     4020
ggcaggactg actgctgtgg tcagaggcgc ccctggggct ttgggagcca cagaatcttc     4080
ctgagggcag cgccggagga ggccccagtg agagtgccca ctgccaggct cattcctcag     4140
gctgccgcag gcctctcccc aaaacaggca atgcttctca gcaacctgcc ccaggagcag     4200
gccagggaag gccgccatcg gcctacagtg ctgggctctg gagggcttgg ttggtaacag     4260
gccatggttt ctatgagcca gctggggtgt gaaggacaca ggctggattc acctctctgg     4320
gcctcagttt ctgcattcaa aaagtgggaa tcatgatatc tgctctattt cttatctctc     4380
agtgctgatg tgaacctcca ataagacttt taaaaatact ctttctacct tacttttatt     4440
tttcatttat tttaagataa tgtctagctg tctcacccag gctggagtgc agtggtgtga     4500
ttacggctca ctacagcctt aacctcccag gctcaagtga tcctcctacc acagcctccc     4560
aagtagctgg aactacaggc atgcaccacc gcacctggat aattttttct tttgagacaa     4620
ggtttcactc tgttgcccag gctggagtgc agtggtgcac tcttggctca ctgcagcctc     4680
aacctccctg ggcttaggtg atcctcacac ttcagtctcc caagtagctg ggactacagg     4740
tatgtgccag tacacccagc taatattttt gaaggatggg gtttcactat attgcccagg     4800
ctggtcttga actccagggt ttaagcaatc taccttcctc agcctgccaa agtgctagga     4860
ttataggtat gagccacccc ccggcctata atcctaccac tttaaaaaag cctgtaattt     4920
tagcacttta aaaaattttt ctaaattttt tatagagatg ggggacagct gtggtctcac     4980
tgtgttgccc aggctggtct tgaactccta ggatcaagcc atcctcctgg cctggcctcc     5040
caaagtgttg ggattataag cataagcctt accttacctt ttttttttga gttgcagttt     5100
tgttcttgtt gctcaggctg gagtgcaatg gcaagatctt ggctcactgc aacctccacc     5160
tcccgggttc aagcaattct cctgcctcag cctcccgagt agctgggatt acaggcatgc     5220
```

```
gccaccacac ccagctaatt ttgtattttt agtagagatg gggtttctct atataccttta 5280
attttaaagc actgcattca tgtaaattgt gattaacatg gattcaagag agggagtgag 5340
gatgaatgag ccaggcagtc acctcggctg tcaccctcca cttctctcct ccttctgaca 5400
gtcatcgtcc atccgtttct gcagctgttt gtttgactct cctgatcatt ttgcttgcca 5460
cataacttgc ctcctgggaa agaatgccct gggcaggccc acatgagtag tgaaaaataa 5520
tctgcagtga aaaataaaac taagtagtct ggtccacaga gcagtcttat tttttcactg 5580
cagatgaagg agttgacatt caggcttcat tctcatttat aagtgtttta aagacacata 5640
cagtggattg aacagtggcc ttcaaaaaga tgtatctaca tcctaatccc tgggacctgt 5700
gaatgttaac caagttagga aaagggtctt cccgggtgtc attaagttag agatcttgag 5760
atgaggagct catcgtggat tatccaggtg gaccctgcat ccaaggacaa atggtcctta 5820
gaaaagaaaa gcagaggctg ggcacagtgg ctcaagcctg taatcccagc actttgagag 5880
gccgaggtgg gtggatcacc taaggtcatg agttcgagag cagcctggcc aacatgatga 5940
aatcccatct ctactaaaaa tacaaaaatt agcaaggcat ggtggcgggt gcctataatc 6000
ccagctactc aggaagctga ggcaggagaa tggcttgcac ctgggaggcg gaggttgcag 6060
tgagccaaga tcgcgccact gcactccagc ctgagggaga aaagtgaaac tctgtctcat 6120
aaaagaaaag aaaagcagac agagatctga gacagaagag gagagtgaag gaaaaaaggc 6180
catgtgaaga tgaggcagag gttggagcca tgcagccaca agccaaggaa tacctggagc 6240
cccagaagtt gcaagaggta ggaagaagcc tcccctagag cctccagacg gagcacagcc 6300
ctgccaacac ctccacctca gacttctggc ctccagcact gtgagataat caactgctgt 6360
tgttttaagc caccagattt gtggtaattt gttatggcag ccacaggaaa ctaatacagt 6420
acctaatctt cacaaaccca tcttacagaa aaggaaactg aagtcagaga ggtagtggct 6480
tgtgcagtgt gttaggccat tcttgtatta ctataaagaa atacctgagg ccgggcatgg 6540
tggctcacgc ctgtaatccc agcactttgg gaggccaagg tgagtggatc acttgaggtc 6600
aggagttcaa gaccagcctg gacaacatgg tgaaacccca tttctactga aaatatgaaa 6660
attagccagg catggtggcg tgcatctgta gtcccagcta ctcaggaggc tgaggcagga 6720
gaatcacttg cgcccgggag gaggaggttg tagtgagcca agattgtgcc actgcactcc 6780
agcctgggag acaagagaga aaccctgtct caaaataaat aaaaaacaaa taaacacctg 6840
agactgggta gtttataaag aaaggggtta actggctccc ggttctgcag gctgtacaag 6900
catggtgccg gcatctgctt ggttgctggg aaggcttcag ggagttttac tcatcgtgga 6960
aggcagagcc agagcaggtg catcacacag caaaagcagg agcgagagag agagagagca 7020
gggaggtgtg cacactttta aatgagcaga tctcacgaga actcaccatt gcaaggacag 7080
caccaagcca cgaggggtct gcccccatga cccaaacctc ccactaggcc ccaccccccaa 7140
cattgggaat tacagttcaa catgaggttt ggggggacaa atatccaaac tatatcattc 7200
cacccctggc cccccagatc tcatgttctt ctcacattgc aaaatatagt catgccttcc 7260
cagtagcccc ccaaagtctt aactcatccc agcattaact caaaaatccc attcccaagt 7320
ccaacgtctc atctgaagat gagttccttt cacctacaag actgtaaaaa tgaaaacagt 7380
tatttactgc tgagatacaa tgggggcata ggcattaggt aaacattcct gttccaaaag 7440
ggagaaatcg gtcaaaagaa aggggctata ggccccaagc aagtccaaaa cccagcagag 7500
caatcattca atcttaaagc tccaaaataa cctccttaaa ctccatgtcc catagccagg 7560
gcacactggt gcaaggggca ggctcccaag gccttgggca gctctattcc tgcggctttg 7620
cagaattcag tccccatggc tgctcttaca gattggagat gagggcctgc ggcttttcca 7680
ggtgcagggt gcaagctgct ggtgatctac cattctgggg tgtggatggt ggcggccccg 7740
tcccgcagct ccactaggca ttgtcccagt ggggactcta tgtggggcct ccaaccccac 7800
atttcccctc caatgggaag gctctgcccc tgcagcagcc ttcttcctgg gctcccaggc 7860
tttctcatac atcctctgac atctaggtgg atggtgtcaa gcttccttca ctcttgcact 7920
ctgcacacct acaggcttaa caccacatgg aagctgccaa ggtgtatggc tggaaccctc 7980
tgaagcagca gcctgagctg tgactatggc cctttgagcc aaggctggag ctggaacagt 8040
ctagatgcag gcagggagca gtgtcctgag gctgtgcaga gcagcagggc cctgtgcctg 8100
gacaatgaaa ccattctttc ctcctcatcc tctgggcctg tgatgggagg gttgtggaag 8160
atctctgaaa tgcctttgag gcctttttgc ctctgaggcc tatttcctat tgtctcagtt 8220
attggcagtc ggctcctttt tagttatgca aatcctctag caagaggtta ctccactgcc 8280
ggcttgaact cctctcctga aaaagctttt tctttctttg tcacatggcc aggctgcaaa 8340
ttttccaaac ttttatgctc tgttttacct ttaaatataa cttctaactt taattcattt 8400
atttgctcct gcatttgagc atagggaatt caaagaagct gggccacatc ttgaatgctt 8460
tgctgcttca aaatttatgg ccacgcttgg tggctcacac ctgtaatccc agcactttgg 8520
gaggcctagg tgggcagatc acgagatcag gagatcgaga ccatcctggt caacatggtg 8580
aaacccatct ctactaaaaa tacaaaaaaa ttagcttggt gtggtggcgc agacctgtag 8640
```

tcccagctac tggagaggct gaggcaggag aattacttga acctgggagg cagaggttgc 8700
agtgagccca gatcatgcca ctgcactcca gcctggtgac agaataagat ttgatctcga 8760
aaggaaggaa ggaaggagga agggaagaaa tgtcttcccc ccagatgtcc tgggtcatcc 8820
ctcttatgtt caaacttcaa cagatcccta gggcatgaaa ataatacagc caaattattt 8880
gctaaggcat aacgaaagtg acctttgctc cagttcccaa taagttcctc atttccatct 8940
gagactcatc accctggcct tggcttgtcc atatcactgt cagcattttg gtcacaatca 9000
tttaaccagc taatcgggag gctgaggcaa gaggatcact tgaacccagg aggttgaggc 9060
tgcagtgagc tgtgatcaca tcactgcagt ccagcttggg caacagagca agatcctgtc 9120
tcaataaata aataaataaa tacataaata acttaagttt atttaaagct gcatctttgc 9180
caccatggag aaaggccagg ccagctcctt ctctctttct gcacgtgttc ctcccacctc 9240
agctgcctct gctcctcaag gaggaacaga gggagtagga aaggccatcc caggaggccc 9300
agcaccccat gacctggctc tggggccttg tgggtttatg gattcccagt gctgagtcat 9360
ccctcacagg ctcttgtggg caccttggac attggtcaga agcatgtggt ccccgggaac 9420
acaccttttc ctgatcatct gggaagggca gcttgtgcca gcgaggccac ctgttcagcg 9480
ccacggcccg ccagacagct gcagccacag ccttgccttt gatcagagca aacaccagac 9540
atgtgtgtca tgcccccaac ccatctccag gggacacatg tcctttcttg ccaggcctga 9600
gatgaacaag agagggacaa gtccccaagc ctctctctcc ttcctgcctc acccactccg 9660
ctgttagatt ctcaaggtgg atggtgggct aactagggca accgaccatc ctggtttacc 9720
tagaactgag gggcattttt caggaataaa actgcaaaag tctggagcaa acaggagcaa 9780
gttggtcact ctggggctgg tggagtcagg tttccttctg caggccccct ccccgcaagc 9840
atgggtggaa cccaggacag gaacacagag caggccccag gaccgggctt gtcacttaca 9900
agtctttttt tttttttttt ttttgagatg gagtcttgct ctgtcatcag ggctggagta 9960
cagtggtgcc atcttagctc actgcaacct ctgccttctg ggttcaagtg atccccctgc 10020
ctcagcctcc tgagtagctg ggactacagg tggcaccacc acgcccagct aatttttttgt 10080
atttctagta gagatgagat ggccaggctg gtcttgaact cctgacctca agtgatctgc 10140
ccgccttggc ctcccaaagt gctgggatta caggtgtgag ccactgtgcc tggcccccact 10200
cacaagtctt aaaccatgcc tcagcacatc aatgccattt acaaaaaggt agagggattt 10260
tccaggcaaa aatagatgaa agacatagga tgattgatca tgtcctgctt aaacataggt 10320
ctgatgctat taagaattga gggctgggag cggtggctca cgcctgtaat cccagcactt 10380
tgggaggccg aggcgggcgg atcacgaggt caggagatcg agaccatcct ggctaacacg 10440
gtgaaacccc atctctacta aaaatacaaa aaatggccgc gcgcggtgac tcacgcctgt 10500
aatcccagca ctttgggagg ccaaggcggg cggatcacga ggtcaggaga tcgagaccat 10560
cctggctaac acagtgaagc cccgtctcta ctaaaaaata caaaaaaaat tagccaggca 10620
tggtggcggg cgcctgtagt cccagcaact tgggaggctg aggcaggaga agaatggtgt 10680
gaacctggga ggtggagctt ccagtgagcc gagatcacac cactgcactc cagcctgggc 10740
gacagagtga aactccatct caaaaaaaaa ataaataaat aaataagaat tgttagtatt 10800
ttgcaggtgt gacaaatgat tctgtttctg tggcagaatg ttctcaggag atctcttttg 10860
aactctcatg gaaagcatca tgctgttggc aacatcacat ttatttttat ttatttatta 10920
ttttttagag acagggtctt gctctgttgc ccaggctgga gtgcagtggc acaatcacag 10980
ctcactgcag cctcaacctc ctgggctcaa gcaatcctcc tgcctcagcc tcccaaagta 11040
gctgggacca caggcgtgag ccactgcact cagcccaatg taccttcaat atttacattt 11100
ctggcaaagg tagcaaaacc ttaacaaatt ttgaatctag ataataaaat tatgaggctg 11160
ggtgcagtgg ccctgacagg gatggctcac atctgtaatc tcaacatttt gggaggccaa 11220
ggtaggcgga tcacctgagg ccaggagttt gagaccagcc tggccaacat ggtgtaaccc 11280
tgtctctaac aaaaatacaa aaaaattagc cagacgtggt ggtgcacgtc tgtcatccca 11340
gctactaggg aggctgaggc aggagaattg cttgaacccg agaggcagag gttgtgatga 11400
gccgagatcg cgtcattgca ctccagcctg ggcaaaagca agagcgaaac tctctctcca 11460
aaaaataaaa aaaaaataaa ttaatgaatt aattaaaata aaataaaata atggatagtc 11520
actgtaaaga aaaaataaat gtatatatca gccaacaagt gatggaatag agcaccccat 11580
ctccctggct ggacagatac atcccacaac acctggaagg cggctccatg tagaactttc 11640
tggactgctt gaggtgctgt gctggagcac ggtgacagag gagctggacc atggacctcc 11700
cccggccccc accaagggcg aggtcccccct gtggtgggtc tgagggaggc atccgtatgg 11760
cctctgcggc ttgggcaggg aatttggggt ccaagtactt ggtgcaaagc ctggaaagag 11820
ggtttgggtg ctgagggcat atcccctggg ccacatgggg gcagaagtgg ggccccctga 11880
agcttggagt cctgggcagg ggcatctatt ttgctgtctg aggccttcag tacttgaagc 11940
aaaatggagg cagaatgtcc caccttaatg cccctgattc ctccaaacca attccagaga 12000
cagcaagggc cagaacaggg atggccctgc ccagggtcat gcancgagga agtggccagg 12060

ctgggatctg aacccaggct aatcccctcc cttgtcctcc tccaggccct cacccctgca 12120
tagagccctc cagctcactc atcctcggcc agctccatct cctcagcttg taaacccccc 12180
cgggattttc ctttcttaaa aaacaaaggc ttggccaggc acggtggctc acgcctgtac 12240
tttgggggtg gctcccagca ctttgggagg ccaaggtggg cggatcatga ggtcaagaga 12300
ttgagaccat tctggccagc atggtgaaac cctgtattta ctaaaaaaaa aaaaattaac 12360
tgggcatggt ggctagctac ttaggaggct gaggcaggag aatcgcttga acctgggaga 12420
aagaggttgc agtgagccaa gatcgcgcca ctccacttta acctggcaac agaacaagat 12480
tccgtttcna aaaacaaaca aacaaacaaa taaacaaaaa aaggcggagc gcgatggctc 12540
gcgcctgcaa tcccagcact ttgggaggct gaggcgggcg gatcacttga ggttaggagt 12600
ttgagaccag cttggccaac atggtgaaac cccatttcca ctaaaagtac aaaaatcagc 12660
caggtgtggt ggtgggtgcc tgtaatccca gctactcagg aggctgaggc aggagaatcg 12720
cttgaaccca tgacctggag gctacagtga gctgagattg cgccactgta ctccagcttg 12780
ggcaacaaga tttgtttctc taaaaaaaaa aaaaaaaaga ctggcccttc cccttcagct 12840
cttcctcagg gtccctgagc actctacacc cccgtctaca ctgagcactc caccctgctg 12900
tctacactga gcactccacc ctgccatcta cactgaggac tccaccccac tgtctacact 12960
ggctgcctcc cgccctcacc tcctgctaag gccattcccc gctgcatctg tcttctagat 13020
tctgcagcct tcagcacgct gggcccctcc tttgtcccct tgagccacct ccagcctccc 13080
cctgagctgc tactcctctc ccagcagcct ccacccaagc ccctccagtc cccaagctgt 13140
cccttgcatc cagcactgcc cttccacgtg cccccttccct ccagcttcac agcagggtgg 13200
ggcctccagg ccctgcccac tgtgcccatc cacaagttgt ggtgggagct ccgaggggag 13260
gcaggggtgt gcatggactt gggacgtcca agtctgggac caggggcagc tggttggtgg 13320
agtgtggagg gggataggga ctttcaggta gagaggctgt aggggcaaga tcgggacggc 13380
ggatgtccct aaggagggct ctgacctggg aaatattgtg cagcttcctc tttgccattc 13440
ctggagctca gacactggcc ggctctcacc ccgcccttcc tgcaggacac agctccatcc 13500
cagtgagttc ctagtgtaga catctccagc agcacggatg ggaaaggaag tcatcaaagg 13560
tgcccaggac cggaggcttt ttctggaggt ggcagaggag ggtgtgggtc tcagggctct 13620
ggctgagggc aagcgtggga ggtcttaggt ctgcaccagc cccgtgaagg cccctcctgc 13680
tccctggtgg agtcctagag ggaacagcag cccctaggct ctagcaggag tgggtagggg 13740
cttttctggc ttcctactgt gccagcagga tagctgggcc tggcactgag cccaaagatc 13800
acatgccggg gcattggcgc agtgaggaac agacccttgc caaagctggc aaagaagacc 13860
ccatggggtg cagctggtga agctgagagc tcaatgtttg gggggagcctg gcaaaagggg 13920
tcctcccctc cctctgcagg ccaggatcgc aggttttccc tacatgttgg taattctcaa 13980
acaatcccat ggccactgga gcaaagatca cagtgggcgg cggcctcggg agcagtggac 14040
agggcacgca gtgcctttga tgccagagcc ctcgccccaa agtcaacaaa ctctgcagcg 14100
gactttgcac ccggactttg ttttcaccat acaaggaaag ggacagatca caggccctct 14160
cgctgccctc gctgagccgg aagctgcagc gtgagctctc tcaagcccca tttctaggtt 14220
ccccaggcgc acccctgagc ccctactcgc ctattaagtt ctcctaatag cccttcaagg 14280
tcttaatgta tgtccattag acagagggga aaactgaggc gagggcaagt gacttgaccg 14340
aggttcctcg gcgagcaggg cgtggagctg agaacctcgt tattactgct ccccacacaa 14400
ccctctggcc gttcttggaa gaaggctgag ccccgggggg gccagagtga cccaaacacc 14460
atgggccgcc tgcggtaaca cgtgcggcca cgaaggggca gcagtttccc gcccggccgg 14520
gctctctccg gcgctcagta tccgtcccag gccaagaaga agaaactcgg ggaggaggc 14580
ggaggggct gcgtgggagg gcgtggaaga tggacgtggc caggggagtg gcagctgcac 14640
acagtggatg ctgttaagat gaagggaaag aacgtgggct ccgagatcac tggacacggt 14700
tccacctttc ttcccgctca ctgcatggcc ctgggcgggt tgttgaaccc ttggaaacct 14760
gtttttcctt ttttcctttt tttttgagac agggtcttgc tctgtggccc agactggagt 14820
gccgtggcac gatcttggct cactgctgcc tcccaggttc aagtgatcct cccagctcag 14880
cctcctgcgt agctgggacc ccaggtatgt gtcaccacag ccggctaatt tttgtatttt 14940
tttgtagaga cgggatttcg ccgtattgcc caggctggtc tcaaactcct gagttcaccg 15000
gatcttcctg cctcagcctc ccaaagtgct gggattactg gcatgagcca ccgcacccag 15060
cagagacctc agttttctaa cctgtgccag caggaataat gatagctgcc tagcttggct 15120
gtgctgggaa ttaagtaaga tgaccgggta gcaaatatga agtattactg gacacagagg 15180
gccccaggct gggttagcag cggtggtcag ggctgctgct tcctggcctg agctcgaagg 15240
agggccctca ttaccacctg ggtgagtcct cgtccaagcc tggcactgct gcgtgggaat 15300
aacttctgcc acccaagttg gcagattgtg tgcaaagtta agtcctgact ctgtggggtg 15360
gacttcgagg cctcttcatc ggacctgctt ccggtgactg cattcgcacc tcctcctgtt 15420
cctggtttaa cacagcccag ctttcctcct gctgagccct ccctgggcct gctgtcaccc 15480

tcgtgccgct gtgcctcgca gtgccactcc ctgtaccctg aatactttgc cctgcctctc 15540
cacccagctg agagtcaggg cccctgtgag gctctgccca gcccgtcctc cgggtttctg 15600
cctctgctga gcacttccct gcatgattgc ttctgagagt ccccccagcc tgtgagcttc 15660
tcaggactgg gacagcttct caggaccgag gcttcctggt ctgcttgcaa ttttacaggc 15720
gggcacattt tcccttggcc aacatcagag actggacatc tgcagatctg tgctagccac 15780
tgagcaccca ggcaccccag caggtagctc tgtaaccaac ccattctgta aagctgaggc 15840
tcagagaggt gaagcgcctg gcctggggcc acagcctgcg tcagctgcag agccaggagc 15900
tgagatatgc acctgcggct ctgctcacag ggtcctgcac agactgctgc tggagccacc 15960
tatgtagagt caagagagtt catgttaact ccctctcaca tccctcagcc agggtggggg 16020
ctgacgatag acactcaggg atggcctacc ctccccaaca acccccgtca ggtttgccgg 16080
atctccttgg aagaaaagtt ctgggcagaa ttccaccgtt ggcctggcct acactctcct 16140
tagtggctta ggaccctcag cggtggataa gttgtgggca gaagagatgc aatcaggatt 16200
ctcacccact cacccccttgc cagccccaat aagctcaata agctgggctc ggtctgagga 16260
agtgtccagg aaatgtgcaa atggcctggg acagccctgt gttcctttca gtaaggttgc 16320
tgaaggtgag gctgaaagtt ggagaaacag aagccagtgc ttatggtttt aattaagata 16380
atggaatgta tgtatgtatg tatgtatgta tgtatgtatt tatgtattta tctttagaga 16440
tagagtctca ctctgttgcc caggctggaa tgcggtgaca caatcatagc tccttgcagc 16500
ctcgacttcc tatgcccaaa tgatcctcct acctcagcct cctgagtagc tgggactaca 16560
gacacacgcc aactatgcct agctaatttt tatttctatt ttttgtggag actgggttct 16620
cactttgttg cccaggctgg tcttgaaccc ctagcttcaa gcaatcctcc tgcctcagcc 16680
tcccaaagtg gagggattac aggtgtgagc caccacacct ggcctggaat ttatttgtat 16740
tctgcttata aaattaatac attcttattg cagaaaagtt tgaaaataaa agaaaggaca 16800
aagaacaaaa agcgtatata atttcacagc tcagatctca ctgctattaa catttttatt 16860
tactttcagg ctttttcttt tctaggtaca tatgcagaga ttattttatt ttatttattt 16920
tattttatat tttattttat atttttatt tcattatttt attttatttt attttattat 16980
ttttagagac agggcctcac tctgtcaccc aggctggagt acaatggagt gatcatagct 17040
cactgcagcc tcaaacacct gggctcaagc aatcccccca ctcagccttc tgagtagttg 17100
ggactaaagt gtgagtctgg ctaatttttt ttactttttg tattgacaga ggtctcacta 17160
tgttgcccag gctgatctca aactcctggg ttcaagcgat cctcccacct tggactccca 17220
aagtgctggg attacaggca tgagccacca tgcctggcct aaaatgccac tttttgtcat 17280
ttactaaaat cccatggaca ctttgacatg tctgtattct atgctattga tctgactgtt 17340
ggcatctaca tcattatggc catctatcat ctatcataat ccattttaac attaaaattg 17400
tgctgctgct tagatttttc tggcctgtct cctatttgta ttcttccaga taaattttag 17460
aatcatttta tcaaattccc cttgcagaaa aagccctatt ggattttggt tgaaaaatac 17520
tgaattttta cattaactta ggaaagggct gggcacggtg gctcacgcct gtaatcccta 17580
cacttttcga ggccaaggca ggtggatcac ttgaggttgg gagtttgaga ccagcctggc 17640
caacatggtg aaactcggtc tttactaaaa atacaaaaat tgccaggcgc attggctcac 17700
ctgtaatccc agcactttgg gaggccgagg tgggtggatc acgaggtcag gagatagaga 17760
ccatcctggc taacacggtg caaccccgtc tctcctaaaa atacaaaaaa ttagccaggc 17820
gtggtggtgg gcgcctgtgg tctcagctac ttaggaggct gaggcaggag aatggtgtga 17880
acccaggagg cggagcttgc agtgagccaa gatcgcgcca ctgcactcca gcctgggcga 17940
cagagtgaga ctccatctca aaaaaaaata ataataataa tacaaaaatt agccgggggt 18000
cgtggcgtgc acctataatc ccagttactt gggaggctga ggcaggagaa tcgcttgaat 18060
ccaggaggtg gaggttgcaa tgagcagaga tcgtgccact gtactccagc ctgggtgaca 18120
gagtgacact ctgtgaaaaa aaaaaaaaaa ttctgaagga ttgagactct tagactctta 18180
ggtcttccta tccaagagca caatatagct tttcatgtat tcaagccttt ttcaatgcat 18240
caacagaatt ttacagtttt tttcatgata tcctgctatt tcttataaaa tgtattccta 18300
gatattctgc atgttttccg gttgtttgtt aataaatatt tttcatttgt cattatttcc 18360
taattggctg ttatttgtat atatgacatc tgttgaattt tttgattact ttgaaaatgg 18420
ccattctttt gtgttttttt ttaactttct attttgagat aattttgact tacagaagat 18480
ttgcaaaaat agtacagaga gttcctgttt ccccccttatg ttaacccagt ttctccttat 18540
gttaacatct tacataacta cagaacaatt gtcaaatcta agaatcaacc tgggcacaat 18600
gctattaact aaactgcaga agctgttcag atctcaccag ttcttctact gctccccttt 18660
tctcttccag tgttcaatcc ggaatcctac attatattta gttgtcattt ctctttggtg 18720
tcttccaatc tgtgacagtt cctcagtctt tctttgtctt tcatgacttt cattttttta 18780
tacttttgaa aaatactggc cggttgtttt gtagaacgcc ctcagtttgg gtttgcctga 18840
agttttttgt gattagatcg aggtcatgca ttattggaga gggtgccacc gcctcgatgt 18900

```
gcaagctcaa tgcatcatat cagagggttt gtaatgtcag tttataccgc cggagaccct 18960
aacctggagc atttcgtgaa ggtgctgtct gccaggattc tccactagaa agttactatt 19020
tttccctttt taattactga atgtctgagg ggaaatactt tgagactatg caaatatcct 19080
gtttctgctt taacttcggc tcactaagtt tagcattcat ctatggatct cgcttatagc 19140
aagtattact gtggagttct aatggtaatt ttctgtttct ctcattcctt caacctttat 19200
taatatgctt cttcctcact tattcatttt gtttcagttg tttataccaa catggatttg 19260
tggatattgg ttttattctt tgggttgcaa ttgaatccta tcattatttt gttagtcagt 19320
tgttccatcc gaccttggtc attaggagcc cttgaaattt ggctcccatg cctttttttt 19380
tttttttgag accgagtctc actctgtcac ccaggtttga gtgcagtggc atgatcttgg 19440
cttcctgcaa cctccgcctc ccaggttcaa gcaattctcc tgcctcagcc tcctgagtag 19500
ctggtattat aggcgctcca ccaccttgcc cggctaattt tttgtatttt tagtagagat 19560
ggggttttat tatgttggcc aggctggtct caaactcctg acctcaggtg atctgcccgc 19620
ctcggcctcc caaagtgctg ggactacagg cgtgagccac cacacctggc ctcctatgcc 19680
attttaacat gcccgtcttt tctttttctt tcctactttc tgtgactgta agaagctcca 19740
ggatacattt ttgctgccct agacttagcc tcaatcagtt ctcagaaaag ctctggttct 19800
ttttatggga tacttagaaa actagctctg tatggcctgg cgcggtggct cacgcctgta 19860
atcccagtac tttgggaggc cgaggtgggc agatcacaga tcacgaagtc aggagatcaa 19920
gaccatcctg gctaacatgg tgaaactctg tctctactaa acatacaaaa aattagtcca 19980
ggcgcggtgg cgggcgcctg tagtcccagc tactcaggag gctgaggcag gagaacggca 20040
tgaacccggg aggcggagct tgcagtgagc cgagatcggc agccactgca ctccagcctg 20100
ggccacagag cgagactccg tctcaaaaaa aaaaaaagga aaaagaaaaa agaaaactag 20160
ctctgtatgc tagttttttt tttaagacag ggtctctctt gccccagctg gagtgtagca 20220
gcacgatcac agctcactgt agcctcaacc ttctgggctc aagcaatcct cctgcctcag 20280
tctcctaagt agctgggtct acaggcatgc accaccgtac gtggcaattt ttaaaaactg 20340
tttgtagaga tggagtctcc ctatgttgcc tggtctggaa ctcctggcct caagtgatcc 20400
tcctgcctcg gcctcccaaa gtgctgagat tacaggcatg agccactgta cctggcctgg 20460
ccaaggtctg tcttttttta aaagaagttg ttgtatagtt gttttttttt ttattttttt 20520
ttctgagacg gagtctcgct ctgtcgccca ggctggagtg cagtggtgcg atctcggctc 20580
actgcaagct ccgcctccca ggttcacgcc attctcctgc ctcagcctcc cgagtagctg 20640
ggcctacagg cgcccgctac cacgcccggc taattttttg catttttagt agagacgggg 20700
tttcaccgtg ttagccagga tggtctcgat ctcctgacct cgtgatccgc ccgcctcggc 20760
ctcccaaagt gctgggatta caggcgtgag ccaccgcgcc cggcctgttg tatagttttt 20820
atctcgagtt ttctagcgat ttaatcatat tggttacaaa aaaggatgat tttactacct 20880
cctttccaat gtttctacat attttttcat tttatctaac tgcattttaa aataaacttt 20940
taattttaga atggtttcat atttacagaa aatgtgcaaa gatagtacag agagttcctg 21000
tgtactccac acccggtttc cttattatta tcttaacgtg atacacaatt aataaaccag 21060
taacattatt attcactgaa gtccacactt tctttttttt tttttctgag acggagtcta 21120
cttctgtcac ccaggctgga gtgcagtggc gcaatctcgg ctcactgcaa cctccacctc 21180
ctgggttcag gcaattctgt ggctcagcat cccaagtagc tgggaataca ggtgcccgcc 21240
accacgcccg gctaattttt tgtattttta gtagagatgg ggtttcacca tgttagccag 21300
gatggtcttg aactcctgac ctcgtgatct gcctgcctca gcctcccaaa gtgctgggat 21360
tacaggcgtg agccaccgcg cccggcgtcc atactttctt tagatatcct tcctttttac 21420
ctaacgtcct tcttctggtt caggatccca tccagaaagc aacattaccc ctcgccatca 21480
cgtcttcaca ggctcccctt gacgggaaga gttcctcaga ctttccttgt ttttgttgac 21540
cttgacagtt ttgaggagga ctggtatctt agtctgtttt gtgctgctat cacagactag 21600
ctgagaccga tacatgatac atgaaaaaaa atgtattctt acagttgtgg aggctgggaa 21660
gttcaagacg aagttgctgg ttggtttggt ctctggtttc aagatggcgc cttgctgctg 21720
catcctctgg agaagaagaa tgcggtgtcc tctcactgca gaagatggaa gcgctaaaag 21780
gaatgaactc cctttgccaa gccattttat aatgggcatt aatccacaaa ggatgaaacc 21840
ctgagaaaca tcaagcttta aagcactggt tctcaacctt tttggtctca ggagcccttt 21900
atactcttaa aacgttttga ggatcccaaa aaaaggcttc tacaggttcc atcttttaat 21960
atttaccata tcaaaaatta aactgaaaaa attttaaatt atttattcat ttaaaataac 22020
aaggataaac ccattacatg ctaacataaa tcatgtattt tatgaaaaat agctatattt 22080
atcaaaacaa aaattagtga gaagagtggc atgtataatt tttttgttt atttttttgtt 22140
tttagatgga atcttattct gtcgcccagg ctggagtgca gtggtgtgat ctcggctcac 22200
tgcaagctct gcctcccagg ttcacaccat tctcctgcct cagcctcctg agtagctggg 22260
actgcaggtg cctgccacca cgcccggcta atttttgta tttttagtag agatggagtt 22320
```

tcaccgtgtt agccaggatg gtcttgatct cctgaccttg tgatccaccc gcctcagcct 22380
cccaaagtgc tgggattaca ggcttgagcc actgcgtctg gcctaaattt ttgtgaatgt 22440
ctttaatgcc tgccttctca tatttgtttc tgcattcaag ttattgcaaa atgttgtgtt 22500
ggttgaagtt tgtaaagaaa atgtggcctc atacagttgt gtagttggaa aggcaagagt 22560
attttgattc tctcttcaaa caactatgga caacctgctg ttacaaaacc agaatgcaaa 22620
aagttgtagt aaatacaggt taggtgtagt gtggaatctg aaagcatgtg aatgaacttt 22680
ctgagttttg taacattaaa gtccagttgc gttaagctac tgtgatagca tatagcattg 22740
tcctaatact ggaattagta tcagaagtgg ggtgctactg ttaataaata aaaagaaata 22800
aataaatcat gtgatactgg ctcagaagtc aggcagtagg ctgtgtggaa cctgacatca 22860
cgccatgtaa tacattggca accatttgat ccagctgtct gtcatgatga cttggaaagt 22920
caaccacata cttacagagc ctgtagacat aggggaaaat agtataaaac agaatactaa 22980
cagtggacct tggttcttgc cagttgcatt tagccaaata ttaaacaaaa gagatattct 23040
tgggcagcaa ctggaccatc ttcaagtaaa agtgaaaggt aataaacaga gtccagacat 23100
ttgtgcccat gcgggttaag aaaaatccag ttgcttctag acaccgtata tgaaaacaac 23160
gctgaaaaca agcctttgag tggtaaaggc cgattaacac tcagcgcggt aacaaagacc 23220
aggtgggcta acccgaaatg aaatgagaag cctgtggtga tgaggaggca gagaagtaaa 23280
atcaagtttg agcatttcgt ttaggagagt ttgggctctg attacttgca catgcaaacg 23340
aactggaaac aaacagatca gatgtctacc acttcttcga gggaattgca ttgccaaaga 23400
agtcatgaaa gcagactcta tactgattag gcattaaaac aaaaacaatc tttaggcccc 23460
taaacttgca tgggcaggaa gtgggctgtc aaagctgttc atcctctaag gtggacctag 23520
ttcctagtcc ccagtataca cttcagatgt ggccctggag gacactggac atggaggacc 23580
tcccagagga tgaggctagg gcttcatttc tccaatgacc tcagctgcct ctatttcccc 23640
ttcttcctct ggaagtccta tcatcgttat tattattatt atcatcattt ttattttgag 23700
ataaggtctc gctctgttgc ccaggctgga gtgcagtgac atgatcatgg ctcactgcag 23760
ccctcccagg ctcaagtgat cctcctgcct cagcctcctg agtagctggg agtacaggca 23820
catgccacca tgcttggcta ttttttttttt cagtagagat agggctctca ctatgttgcc 23880
agggctgatc tcaacctcct gggttcaaga gatcctccta cctcagctcc tgagtagctg 23940
ggattcgggt gcacaccacc atgccaacta atttttaatt ttttttgta tggacaggat 24000
gtacagtgtt agaaatggat tgcttgcaga ggcaggagga tcacttgagc ccaggagttt 24060
gatcacactg tgaaccatga tcgcacccct gcactccaat ctgggcaaca gagtgagacc 24120
ttgtctcaaa aaaaaaaaaa aagagagaga gagagagact caaagatagg caaaaaagtg 24180
ggaaagcttt atagtggaca aaaaggaacg ctctaagtct gccctattgg catggtgctg 24240
aaggtgggct aactagagat agggggtact atgtggttga ctatgggtgc atctttggct 24300
ttccctgggt gatcctaagt tggaagcagg gacaaaaatt agggaagctg ttagttattc 24360
atcacgttct ggcagtagtg gactggttgt gatagaagtt attgttttgg ccaggtgcgg 24420
tggctcatgc ctgtaatcct agccctttca gagttcaacg tgggtggatc aggaaggagg 24480
gaggatttgg gaggtcagga gttagcctgg ctaacctggc gaaatcccat ctctactaaa 24540
aatacaaaaa ttagctgggc gtggtggtgc atgcctataa tcccagctac tcgggacgct 24600
gaggcaggag aatcagttga acctggggag gcggaggttg cagtgagcca agatcgtgcc 24660
caatttcatc tcaaaaaaaa aaaaaaagtt atcgtttagc ttcctcgatt gttactggac 24720
gtagtaatct ggcttcctgc aagtctaact ttcagcagac tggctacatg ggctgtgtac 24780
tgtagataag gcagtaagta aagcaaaaat tgatagagca tcaaggataa atagaaaatc 24840
cgtaatcaag cagaagattt gaacacttca ctttcagtaa ctgataaaac aagtagacaa 24900
aaaaaatcag taaggatgta gaagatttga acaacgtaat taacaaactt gacttgattt 24960
acacgtctag aaccctgcag aacacacact ttttcaagca tactcagaac atttatataa 25020
agtgaccata tggtggacca taaagcagtt tcaacaaatc tcacaggagt aaaataacag 25080
accgtgtttt ctgaccgtaa gtacagttaa cctagaaatt gaaaacaaaa agctagaaaa 25140
accccatgta tctggaaatt ttaatataca ctttgaaata acaaatggat cagagattaa 25200
ttcaaatagg aatttagaaa taccttgaac tgaaaaataa tgagaatact ataccccaaa 25260
actgtggggt gcagctgaac agtatataga cgaaaagtat actcatatgt gcataccta 25320
aggagcgggg aggattgaaa gttaatggga ggcaaaagca ggtggatcac ttgaggttag 25380
gagttcaaga tcagcctggc taacagggtg aaaccccatc tctactaaaa atacaaaaaa 25440
ttatccaggc gtagtgaggc tgaggcaaga gaatcgttgg aacccaggag gcagaggttg 25500
cagtgagccg cgattgcgcc actgcacccc agcctgggag acagagcgag actccatctc 25560
aagaaagaaa aaaaaaaaag aaaaggccag gcgcggtggc tcatgcctgt aatcccagca 25620
ttttgggagg ccgaggtggg cggatcacga ggtcaggaga tcgagactat cctggctagc 25680
acggtgaaac cccgcctcta ctaaaaatac aaaaaaatta gccaggcgtg gtggcgggtg 25740

cctgtagtcc cagctactca ggaggctgag gcaggagaat gtcatgaacc caggaggcag 25800
agcttgcagt gagccgagat cgcgccactg tactccagcc tgggcaacag agagagactc 25860
tgtctcaaaa aaaaaaaaaa gttaatggga taaacatcca tctcaagaag ttagaaagga 25920
atgacaaata aaccaaaaaa aaaaaaatca aaagaagaaa atcataaggt caagactata 25980
aagagagtgg ctgggtgcag tggctcaggc ctgtaatctc agcattttgg gaagcagagg 26040
tgggcagatc acttgagccc aggagttcaa gaccagcctg agtaacatag agagacctca 26100
tctttgctga aaataaaaat aaaaaattag ccaggcatgg tggtactgag gtgggaggat 26160
cacttgagcc taggaggttg aggctgcagt aagccatgat tgtgccactg cacttcagcc 26220
tgggtgacag agtgggaccc tgtctctaaa aaactaaaat aaggctgggc gcggtggctc 26280
aaatctgtaa tcccaccact ttgggaggcc aaggctgagg tcagcagttt gagaacagct 26340
tggccaacaa gatgaaacct catctctact aaaaatacaa aaaattagtt gggtgtggtg 26400
gcatgtgcct gtaatcccag ctacttagga ggnnnnctnt ngattatatt ttctccttcc 26460
tacgtcgtta ttggactgaa ttcagaatga tgactctcat tggagctctt cctgtctcct 26520
aactacagtg gcttccgacc ccactctggt tttcacttca cccctctgct gctcatacga 26580
gtagatactt ccttccttct ttctcacttg ttgctcttcc tcaacccccc ccgttggtgt 26640
cccctcctct ttatcttttt ctcgcgacac ctgcgttctc ttgccctctt atcatccctt 26700
tctcgaggcg gtcctttcct ttatccagct taaatacctt ctcctctgtt tatttggggg 26760
ttgggttttt atctctcacc ctccctctaa tttctttcct ctttccgcac ccatcaagcc 26820
tctcgtggtt tctcttcctc tactctcggg tccccccccct ctccccttct ttttttcttc 26880
accccccccaa gcgctttgcc tttttttct ttgcccttta ttcccccc 26928

<210> 7
<211> 29430
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> (4336),(4345),(4349),(4392),(4447),(4490)
<223> Identity of nucleotide sequences at the above locations are unknown.

<400> 7
aggggaaggg ccggctccgt agctcacacc tataatccca gcactttccg aggagagagg 60
atcatctcag gccaggagtt caagaccagc ctgggcaaca cagcaagacc gcatctctac 120
aaaaacttct tttaaagctt aaaaaaaaaa aaaaaagcaa agaggacagt tcaggagaaa 180
agcctgtaga ggcagcacac taaggaggag acgcagccca ggcaccagga ggggctggcc 240
atgggcactc actcctccag caggcgagtg cccagcacca gctggcccac ccagacaccc 300
aggacacggc ctgaatggct ccgtattcac gtgggtggta ataaacaagc aatacacata 360
gccaataagg acaccttagt aatgttacat cataaacgct gcagatcagg gaaatggtgc 420
agggtgaagt gggttggggg gctgcatgct acatgagaag tgggtcgggg ggctgcatgc 480
tacctgagac agagcaggcc ttgctgggaa agaaggagcc ggcaggcctg ggcaaaggtc 540
ctggggtggg agcacactgg agcagagtgt gggggtagca tggcgggtgc tggtcctctg 600
ggcgccttcc caccacgtca tgtgcccatg tgcccaaggt ctctcgtttc acagcccccct 660
gaagctcagg ggtcacagct acacagcccc cagatacctt ggcctgcccc aggtcattcc 720
atccagtgat ggacctgctg acctctagcc tgacctctgg gcagcgtaat ttgagaagga 780
ggagaaggga gggcaacaga cctggggcga tgagggatgc acagggtggc agacacctga 840
ggctgcacct tggagcctca gttctgggtg tgggtggggg atggacaggc tgagggctga 900
agcagctggg cccggccacc atcacacccc aggacccacc agatcaccat gaaaaaccga 960
atgtcaactg gcagcccaga gtgcagaaca aacctttcag aaacacggtg gtgactgccg 1020
catcatgaac ataaaataat tacgccctct ccccagggat caccccctgca ggagtttgtc 1080
ccaagaaaca ccagaaagaa ggaaaacgtc tgagtcacaa tatttgctga ggccttattt 1140
gtaatagcaa aaaaaaaaaa aaaaaaagaa caatctccag cggcaggggt aactagacta 1200
ttgtctccgt ggaaaggtag caccaattaa ctagtaacaa aatgactgcg gtaacaacaa 1260
aacgttcgac atgtcaacac caaaaaccac acacccagca taaccgtgaa ccatgatttc 1320

tactagaatg aatggcagtt atgagaaagc accagcggag acaaagattg aaaaagtaaa 1380
ggtggcctca ttagggagac aagtctctgg gtaatatatt gtaatactgg taaatatata 1440
gtttttaata tattttttaa ttccaaattc catatatgtt cctatgaagc tatttctgca 1500
aatatttttt tcaggaccgt acatcacaaa ggcaaaaggg ccaggtcagc tctccagctg 1560
agagtgacca cttcagagca gacggcagac tccagggtta gcaagcctgg ctgagacctg 1620
gcccatgaca atcactcaac ccctctgacc tcaacatcct gtctgtgaaa tggggataat 1680
tactgcacct ccacatcaca gagtgcgagg cttaaacagg atgcttcata gaaaagcgct 1740
caagaggtaa cagccgggag ggggtagtgg ttttcattaa ttaaatgttg ccttcatcca 1800
gccctgggcc agctccaaca caaagcacac accatccact cagactcagt tgcctggatt 1860
caaagcccgg cctggcctcc agctgtgaga ttccgggcag gatttcccat ctcccagagc 1920
ctcagtttcc tcattcatga aacaggaagt gatcattcct tttattttta tttttatttt 1980
tattttgaga cggagtttca ctctagttgc ccaggctgga gtatgatggc gcaatctcag 2040
ctcactgcaa cctcggcctc ccagtttcaa gcgattctcc cacctcagtc tcctgagtag 2100
ctgggattac aggcacacgc caccacgccc agctaatttt gtatttttag tagagacggg 2160
gttttgccat gttggtcagg ctggtctcga actcctgacc tcaggtgatc cgcccgcctt 2220
ggcatcccaa agtgctggga ttacaggtgt gagccaccaa gcccagttga caactgcttt 2280
taaagacacc tctggctgct gtggaaaaca gcctggtagt gcctcaaaaa gttacacata 2340
gaatgatcct atgaccagta attccactcc tacatatata cccaaaagaa ctgaacccct 2400
ctactcatgt atgtacacat acaggtacac gcatgttaac agcagtgttc acaaagccaa 2460
aacatggaaa cagctcaaat gtccataacc gatgaacgga taaatgaaac gtagtctatt 2520
caccacctga cggaggtgag aggggccata aaaaggaatg atgcataaaa acgaatatta 2580
tggccaggta tggtggctca cgcctgtaat cccaggactt tgggaggctg aggcgggcgg 2640
atcacgaggt aaggagttcg agaccagcct ggccaacacg gtgaaacccc atctctacta 2700
aaaatacaca aattagctgg gcatggtgga gggcgcctgt aataccagct actccggagg 2760
ctgaggcaag agaatccctt gaacctggga aacagaggtt gcagtgagct gagattgcac 2820
cactgcactc cagcctgggc gacagaccaa aactccgttt cggaaaaaaa agaaaaaatt 2880
agccaggtgt ggtggcgggt gggtccctgt aatcccagct ctacttggga tactgaggca 2940
ggagaaccac ttgaacccgg gaggtggagg tagcggtgag ctgagattgt gccactgcgc 3000
tccagcctgt gtgacagaag gagactctgt ctctaaaaaa caaaaacaaa aaaggcccga 3060
cgcggtgtct tacacctgta atgccaacac tttgggaagc caaggcaggc agatcatctg 3120
aggtcaggag tttgagagca gcctgggcaa cacggtgaaa ccccatctct actaaaaata 3180
cagaaattag ccaggtgtgg tggcacatgc ctgtaatccc agctactcgg gaggctgagg 3240
caggagaatc gcttgaaccc aggaagcgga ggttgcagtg agccgacatt gcaccattat 3300
actccagcct gggtgacaga gtgagattct gtctcaaaaa aaaaaaaaaa aaaaaaaaaa 3360
ctaaacaaaa gcaaaaaaac caatgagtaa tgttgtcaag tgaacttcat cccaatggga 3420
atgcagataa tttgtttaaa aggcaccatg cacactgggc aggctggctt cccctgggaa 3480
cgtcttcttt tgcctggatt cccagttggt ttaatcgggc gtagaacact ttcttcaatc 3540
cgggattcag gcacccctgc tcagcacaaa ctcagtacac cccgcactct gctgtgggtt 3600
cttggcacta ttaggagaat gtgagggggt gattcagatc tatctctagt gggtgcatgt 3660
ctgccactcc caggaacgcc cacttctggc aagtcagtgt cagagaaagg ccagctcgtg 3720
gcccctcctg ccttgagtcc caggacccgt gatcagtcct acccggagca gaatcaggag 3780
tttgaaaacc caagtgccaa caatctcatt ttaacccatg taagcatatc caatatttat 3840
atatagaatt cataacagat gtctgggctt ccattccaat agcctatatt ttacactgtt 3900
tatttacatg gttacaccaa acaagactca attcaaggta acccaatcct ttgctactat 3960
accaaaataa gcaacatttt cagtccatgc cttatatata ttcaccaagc attacactag 4020
gcctccaact gctcatcgga gcaagctgca gcctggacac aagctagaga ttaatcagtc 4080
aggaatgatc ctgcgtccag tgccagcatg atggaagaga cagagaaaca gaagacatca 4140
gggctccaga gtcaaggagc ctgcaggtta gttgggcagg atatacacac atacacacac 4200
acacgcacac acaaaaccac ccaagaagaa aaggtgggat gaatgcatgg acaggtaatg 4260
cctggagcct ggggatggat aagctgactg caggtggccc aggcaggctt cctggaggaa 4320
gaagacctgg ctgtangtgg ggtangcang ctttctaaat ggggaaaatc tggctgtggg 4380
tggagttggc angtttccga aaagaagaaa agctgactat gggtacacct ggctgttggt 4440
ggaacangca ggcttcttgg aagaagaaaa tctggctgtg ggtggatcan gcaagcttct 4500
tggaagaagt aaacctgact atgggtggac caggcaggct tcctagagga agaagaccgg 4560
ctgtgggtga accaggcagg cttcctagac agaggaagat ctggctgcgg ttagagtggg 4620
caggcttcta agaagaggaa gggctgactg tgggtagacc tggctgtggg tagactgggc 4680
aggcttcctg gaggaggaag agctggagca ttgaaaaaca aacatgactt ggtgaatgtt 4740

gagcatgccc aggcctgatc cccagaggca attacgcact caagttactt aattctactc 4800
acaatgcctc acaaacaact tctctgacac ctaacacagc tctgggcacc ttctagcttc 4860
agctcctcaa agcagttatt cacgctacta ccctgcacac ctcctcacac cccaacccca 4920
gggacaggag ttctgccaga tgccaaagct cctgatgcca aagcctgggt ctgcttccgg 4980
gctcctcttg gtctaactgt ccacccccgca tcggcatgat gtgcaaaaac aaggctttgc 5040
aatctgccct gatgcctggc ggagcgagtc cctcccgatt cgtctccttc agaaacacct 5100
gggctgccct ggtcctgtta tacccccaac acattctaca gtcagctccg caagttccac 5160
aaagatcaac gctggcgttt ttatggcatt ttatttacag tttttacaat ataaaaaagg 5220
aaggatgcca cagctcagcc agcaggacag acagagatct atgatgcttc tgctgcacca 5280
ttgtttgtgg tcaagaaagt ctgttttcaa tgatttatta aattgtggtg ggagatggat 5340
ggtggcagtg gttaccagca acatgaatgt tcttaatgcc actgaacttc acacttacaa 5400
atggttacga cgataagtgt tatatgtatt ttaccacaat taaaaacagg taaatgcagg 5460
ccgggcacgg tggctcacga ctgtaatctc agcactttgg gaggccaagg caggcagatc 5520
acctgaggtc aggggttcga gaccagtctc gccaacacgg tgaaactctg tctctattaa 5580
aaatacaaaa attagccaga tgtggtggtg catgcctgta atcccagctt ctcaggaggc 5640
tgaggcagga aaatagcttg aaaccgggag gcagaggttg ccatgagctg agattgtacc 5700
attgcactcc agcctgggtg acaaaagcaa aactctgtct caaaaaaata aaataaaata 5760
aaaataggta aatgcaaaca tatggtatag taatattatg ggctattatg agctacaaaa 5820
aagaatgact tgggactaca gttacagccc tcattcagga atttgtttta aatgtgggtt 5880
ggtcgctaag gcatgtacac aacattttga cgttcaaata ttcctagatt tggacagtga 5940
gcacccctct aagctggctc ttctgtccca gaggtcccca ccagtcctcc agaacttctt 6000
tgctttctta cacaataaga tgccccatgc tcggcttgta cctttccttg ccccagccct 6060
agaaccagct tcttcgtgga caagctctga ctcctttggg tggagaatgg tattcagaaa 6120
cccagacctg ggctctggtg tgctcactgc tacttggggt cattgcttct aggcctctct 6180
gctgatggag gtaggatata cacgtacagt cttccctctt cccagattcc gtacttgagc 6240
tcgcctactt gctaacattt atttatatcc cccaaattaa acctcacagc acttctgcaa 6300
tcactcactg acttgcagag tgtgaaaaaa ctgagtcacc atcacacgtt ccaaactgag 6360
gtcaactgag gccacaacgc cccatcttct tgctccggct gtcgagatgt aagcaagtgt 6420
ccttctctcg gtctagctag tgccatgctt tccacatcac tgtgcttttt gtgggcaatt 6480
ttgctgtata aaatgtcccc tgcacatatg ctgctgtgta gtgctcctag gtgcatgagg 6540
ctgccccacg ccttacagag agaatatgca tgagaggctt tattcaggta tgagttatag 6600
cgtagttggc catgaattca atgttaatga atcaacaata tacagtaaat aaggtgcttt 6660
ttagagacag ggtctcactc tgtcacccag gctttagagt ccagtggtgt gaccttggct 6720
cactgccgcc tcaacctcct gggctcaagt gatcctccca cctcagcctc ccaaactgtt 6780
gggattacag gcgtgagcta ctgcactcag cctaaataag gtgtcttaga aacacacata 6840
agacaaggtt atgggctgag tgcggtggct catgcctgta atcccaacac tttgggaggc 6900
caaggtggga ggttcacttg aggccagaag tttgagacta gcctgggcaa catggcaaga 6960
cctcatctgt atattttttt aaatcagaca ggtgtggtgg tgcatgccta tagtcccagc 7020
tactggagag gctgaggcag gaaaatggcc tgagcccagg aggtcaaggc tgcagtgacc 7080
catgattgta ccactgcatt ccagcctggg gtgacacagc aagacgctgt cttaaaaaaa 7140
aaaaaaaaaa aagccaggtc aggtatcgaa cagttggcaa aaacgttgtg acctgaggct 7200
cacaggaacc tagcccgatg tttcccctag gagcaatggt tcagtattca ataattcagg 7260
gttcccagtg actttatgga gcataacttt caagaataac aagaaccaac tgtacgtgtg 7320
tatgtatact cacactttta ttttatttta ttttattttt tgagacagag tctcactctg 7380
tcacccaggc tggagtaaaa tggcgtgatc tcgactcact gcaacctccg cctcccaggt 7440
tcaagtgatt ctcagcctcc caagtagctg ggattacagg tgtgccccca caaccggcta 7500
atttctgtat ttttagtaga gacggagttt cgccacattg gccacgctgg tctcaaactc 7560
ctaacctcaa gtgatccacc cacctcagcc tcccaaagtg ctggaattac aggcatgagc 7620
tgccgtgcct agcctacata cacttttata cacacatgca tctatgacta tttctctatt 7680
tctgtgcatg tgtgcgtggc agtacctaca gtttcagcta tgtgtctggg tactgtctcg 7740
tccaagtttg taagcacctt ctccaaagtg caaagcctgg cttgtgttac tatccatatg 7800
tttacttatt tgctcaatca atttacttat tagctccata accagcttcc catctgctcc 7860
agtagcctct gctgtcagtc acctctgcac cctaccccac cttgcttccg gatgctggat 7920
gccaatcacc cccgacacct ctacatagca ccaccctcga catgctgctt ctttatttct 7980
tatttatttg tttgagatgg agtcttactc tgttgcccag gctggagtgc agtggcacga 8040
tccaggctca ctgcaacgtc cgcctcctgg gttcaagtga ttctcctgcc tcagcttctc 8100
aaatagctgg gattacaggt gcccaccacc acgcccagct aatttttgta tttttagtag 8160

```
agatggggtt tcaccatgtt ggccaggctg gtctcgaact cctgacctca agtgatccac    8220
cttggcctct caaagtgctg ggattacagg tgtgagccac cgcgcctggt ctgcttcttt    8280
aaatgccagg caccaacatt tgtgcaatgg ggtgggagga aagaacaggg aggagagcac    8340
actgccggcc cctgcactga atccactgat caatctgggg gcaactgcca tctccatctc    8400
ctgtcttcct atccgtgaac atctactgca gtcctctcca atgtccttct gtaaagttgt    8460
attatgtttt gcatacaggc cttgcatatt agttctcaga tataatccat atactttata    8520
taaaattcaa accacattta aaaaaataaa actagcatga ctataacgga gtctgcaaca    8580
ttctcacaga ctttatgata aaacatgaaa cttcaaagat acttagggtg gggcagggac    8640
aatgtttaag gctgcctgga agcctcccca tccctgagcc agaaagtcct atctcccctt    8700
caaggggaaa tgcttgaaaa agcactgatc aggctaaaat gacagggatc agggagtaat    8760
caaagtacaa gtgagctggt ctcctccatt ctgagcacag caaagttcag tctctccaag    8820
tccaagaatc atacacctgt ttgccaagaa tgaagttcag gtgtctacaa gtggctgaaa    8880
atattcattg ctgggccatt aacaacattc ttggcaaaac cataccttag cttctcgtgg    8940
aaatttctta aggtagaaga aacaggaaac acccaggctc gcttttatgt agacagttcc    9000
atgaagccag ggaccttccc cacatccacg tttcaattac ctgcacgcag ctcacagtgt    9060
attcaacatc tacgcgtctc tcctactggg gtggcggtgg ccactcaaac cctcatgcag    9120
ctacgatgac cgcaattttg gcaacataat ttcatgtttt tccttgggct tttacccaag    9180
tcagtgacac aattctgcag ttgtctaaag attcaaaatg agggacttga catttacaac    9240
aataataaaa tcttgggttt cctttaacca agcacatgtt ctgcctttta gagaaagctc    9300
tgcaaactca agctggagtg ggatacttgc tgacatcttc aagcacccca ggaatagctc    9360
tactccccca tttccacctt ggctgaacca tctatatccc accaattccc ccaacatccc    9420
tccatccgtc catccatcca cccaaggacc tgctaagcca ggaggtctct cccatctacc    9480
ccacagcctg gcctcagccc acaagggctc tctctacatg aatcccaccg caccagagta    9540
gaccaagtct cccgtagact ccaccctgac cacctccatg cctccagcca ttcccacccc    9600
taaaaaccct ccctggtctc tacacccagc tgatgaatac ttggctgaat gtgacctggc    9660
ctcctggacc caggtgaagc ccacgtcctc cgtaagcccg ccagctcacc ctgcctctgc    9720
accttcactg gagagagccc gcacttcacc tcctcagggc aggcatggct gatgccaccc    9780
agtggaatct ggtgcaaagc agggcccggt gcagagcagg gctgcctgca gagcaaggcc    9840
ctggtgctgg ggccgagcac ctccaatgct ggccgtggaa ccatccctcc cattccaggt    9900
gctgtctcca tcaagaatga gcgagctgct gacatttgca tgacaataat gaataaatac    9960
catattttgc ttcaaatcca gaatagatgt ggccagggtt ggcatatgac tgttgggaaa    10020
ggacagtttg cctcttccca aaccaacttg gattataaaa agcttttctt aacgaccaca    10080
agagcggagg agctcagggg cagacaaaag gaaggctggc tgcagaaggc gggagagtgg    10140
ggccttcagg ggcgggtggg gagagagaaa gcctggagct gcacccccaa ggtctgtgta    10200
catcaggtgc tacagaataa caccacctct tccagcttgg cccccacctg ccctctccca    10260
gcccagtcac ccagacagca ccccactccc cacacacacc tcacatctgc ccgcctcaca    10320
ctcaccagct tcggctctca atgcaacctg gaacctgccc ttggcctctc agctcagcca    10380
cccccattcc tgttggcccc tggcccccca tcgaattctc tctaatccta atgcacacac    10440
ttgcacactc aaacacacac acacacacac acacacacag cccagaggaa aaccataatt    10500
gactgaggtc caggcaagtt tcccgagcag ggaccacatt tcaaaggtca gggaagcagg    10560
cgaacaggaa acatacaggg ggcacgtttg ggggtggagc aggaaataag aaatcacttg    10620
caaaagataa aaagaaaatg aggtagctgg tttcagacac ctcggagcac acagaacagg    10680
acaggcgcct ccgggtcttc cctcaacagg gagatgggcc aggcaggtcc ctgctgctcc    10740
accgcagagc tgggggctat ggccctgaca ccaaggccct ggggcaggcg gggaggcagc    10800
tgttctcctg cctgtgctcc cgggcagggc ctggccccac aagggaactg gccgaaggct    10860
ctgcttggct actccggaaa gtcctgggag acaagcaaag gacttgctag gtcactccaa    10920
acggcccaga tgtgacaact gtgaagaagc cacaccaaag caaggtgaca gaacaatgtt    10980
ggtgacgtca ggttatcagc ttacgctcaa ctccacttac ccggactcac ccgtaacctg    11040
ccgtctcttc ccaaccagta aaggatgcct aggtagaggg gcacaaggcc tggagcataa    11100
ttaccatttt aaaggctctg agaagtcctg cggtgaggaa gcctagttca ctttctctcc    11160
cctaggattt cccaactgcg cctgatcaca gaacattttt tcatttccac tcaggaaaca    11220
tattttgaaa aacactggcc tagaggcaga agtgaaatgg aaaacacaaa agtaaaactg    11280
aacaggaggc actgggcaga gaacggtcag aggcgccctg aatcctggac cggtggagat    11340
ccccagcttg gcatgctccc ctccctgggc ccagaccgcc tccccccatt tcctggataa    11400
gaaggctaat gcgcatcagg gtgaagggct tgcctgggct acacccccag gctcgcccca    11460
caccaatcgc gctcctgcga gagccagtga ctttcttgat ttggctactg tggaattgtt    11520
tgcaactaac caccccagat acagatacaa atgacaggat gatcagatgt aaaggaccca    11580
```

caggtctctg tgatacggct tcatgcagcc agcatggcta gtgccgtgca gaatgagaat 11640
gaccccaggc aagtccttgc ctcccagacc cagaacccca tggagcccac cagggctggt 11700
tcacaagcac tgtctgggtc gggcagagat tccagcaaga ggagggaaca tccatgcacc 11760
ggagccagtt accagaagca aatcgcctct tccaaaaccc aggctattaa tggagtccac 11820
tgttgagtgg agctggggtc tagctatgga atactgcaca gcagagatct tcctgagaga 11880
aagcagtttt ccctgaaagc catgtgtcct ccactaactg tgttttaatt gggcgaacgt 11940
ctgtatctca ttgcagtggc cgcgcatgtg ctgacaaggg gctgggggcg gggtggggag 12000
cagaagctca ggggcctggg agggaaggaa acaggccacc agggctcccc agaaggcatg 12060
tatctctctc acaaacacac gcatgcacac acacgtgcac acatactctg caagccctga 12120
gttagcaact gtggaatgtg accagctcag tgatcccagg acaagctgct agggaatatg 12180
acatttgatt gatgtctgca aatgtgcgtt ttcactaatt agaaggttta gggcagagca 12240
gagaaaaata tgtatttcag agtcccagtt tgacctgcca gaaaccagcc cattactaac 12300
attcttattt tcaacaaaat atagcattct gattacatac catcttggtt ccacgcctcc 12360
tgccttgcca agcccccgga agcggcccaa ggccatggca aatagtgaga gaaacagttc 12420
cagggtggag actgactcag ggtgtcagt cagtggggcg ctgatggccg gtgggaggcc 12480
agcagtcatc accctctcct tgggacagtt gagtagctct cccccagggt catgtggcca 12540
ctcaggttca tatgggaggc gagaggagtg gcagagtcca ggagagtggc tccgaagtca 12600
ctgttccctc caggcctcag tgtcttcatc cattaaatgg gtaggctgag gtctgggatg 12660
acaaggaggg cttgcactta ctgaaaccca tgggaggctg ttcgccgatt tcttttattg 12720
atggaagaaa acactcgtat aattcaagta ccaattaaaa ggcaggcact ggaaccaccg 12780
tctgccaatt cctagttttg cctataccaa atttgagcaa gttaattgac ctctcccagc 12840
ctcagtttct tcgtctgtaa aatgagggta gggatggccc ccagcccaca gggcagctgg 12900
aaggattaaa gaaatcaaac atctcttaga gcccacctgg cacactgtga tacacaacaa 12960
atgttagcta tttttgtcta tgaagtctag attttatatc ttgggtgttc taaagcagga 13020
tacatttatt taaaaacaag gattttcatt aaacacgtac cccacagaca gcaaccccat 13080
ggagactgct cttaattcag gccagtatcg aaacgactct aactacaagc tttatacagg 13140
tctcttggct gtccttcaaa tccaactaag gtggtacttc tgaagcactg tgcacatgtg 13200
tgtgtgcatg cacacgtgtg ggaagggcgg gctcacggat ccctcaggta ccccacccac 13260
gcagtctcaa gtcacaaagc gacagagcag ccgaggaagg tctgtgcccc actggaccct 13320
cgtgaagcca ccaactctac ctctgcgccg tgtcctgcag actgggctac cctttgggtg 13380
gggaccagca tttgatgcaa gaaaggcaga cagaaaagga aaagggcaag ttcgactcca 13440
gataacacag acagtaccaa gccccagggt ccataaatgc cacgcagatg gaagcattta 13500
ctgcgaggcc acacagcaaa cgcacggatc cagggacgga ggtgcagact gcggtgcccc 13560
tgagccatga ccctgcaaat taccaccatg ggaaaggagg ctgccaaacc ccccgacagt 13620
cggctgggct ggcacagact cgtggtttcc atcgaggtgg gaggaggtgg gacgtcccag 13680
cccctccccc atgcccactg cagagggaag cggccgtttc ccctgtgtgg ttacaaaggt 13740
ctcattgttc ttcctcacag ggaggaaact ggaggaccga gctcagaacg cattttagaa 13800
ctggcagaaa agaacatctg gggaaggaaa cacatttcag aaacaaacat acctttgtac 13860
cagcttttat tttctttaag tgttgaaaaa ataataataa taaagacatg ccaaatttat 13920
catcgctcta caaaatccct ttattgagca aaacgtggca gctctacttt caaatgatta 13980
ctgttcctgg aaaattgcag caacgtggat gccaaggccc gaaggccgcc atcagcagcc 14040
aaacaaaaga tgccacctcg ggctccgcga cactgtacca tgccagggaa ctggacagat 14100
ttggggaatg ccacggtttg cctttaaccc cttgcctcct ggtctcctga tgcatctcag 14160
aggctaacat tctttgagga actggcattt cttagttgta aatatgcatg tgggtttggg 14220
agctgcctgc aaagtccagt gttgacgatc agctttgatt tccttggaat caagtttacg 14280
tgtcgagtct ggaagttaag aagaatttgg agaagctgag cactatggtg ttgcaggccc 14340
tgggtgaact cttccaccaa gcattcattg tggactgaca gcgtgcgagg ggctctgcag 14400
gcaggtgcac aggacgaaac acattccgtc cggggaaac ctgcaggaaa gctccctctt 14460
cttcctaagg tgccgggcct agcttcatgg gtccctaccc tccacgcctg tcacactttc 14520
tgagtctcat gtgggagctg cttctggttc ctgacttcac tcagtcctca taggaggtgg 14580
aactactgtc accccatttt acagatgggg agactgggca caaggggacc aagaaaccaa 14640
tgcaaagtca cacttgtggg atcagtgaca ggggagatca attcccaggt tctttctgca 14700
agagttaaat tgttttcatg ctgcctaagg gggggcaact gaaagaccac tgcatatctt 14760
tgccaaaagg gtcaagcaca ggagccgcag ccagtgggtc agatccgcag aggcgctggg 14820
gtgaccctcc ccatacctgg agggatgctt gtcccctcct ggccttcact gggtcccctc 14880
atgaccgtgg cctcccagga cctcagcaca atcccggtcc tgtgctccag gacaagccct 14940
ccgtccccaa gactgtgagg aaatggaacg aagaggggct cgctgcagcc cagcacccac 15000

```
actgcccctt ctcaggggca agaaccgtcc tggaggactt ggctttggag ggggagcctg     15060
ggaggccagt aagtcaacaa gcctctactg ctcatgggtg ggatcccacc gcaggccccc     15120
acctgctggg gcgggcaggg acgggcggca cagcttggcc agggcagata acccccacct     15180
tggccagggc gaaggcagga cacgtgggct ccagcctggc cccaccatcc ctgcacaaca     15240
ctgggcaaag tccacgtttt cctcaactgg gtgttgacat ctgcaggaca ggggcatgga     15300
ggtacagagc gctgaagcca cacagcaacc taggagcgag actccatgcc tccccgggga     15360
cccctcccca ccatgaggac catgaaggct tcccatgtgc cgcaaggact ctggtgtgga     15420
gacacacgtc tcctacacag ccaggcctaa cgctcttgta actgggtggt cccacctggg     15480
ctcacagctg gagggccagg agctcaaggc ttcgcagggt ctgctctcat cccagaggcg     15540
atggggagcc acagcaggct gcaggagaga gggtgggccc cctccacttc agaggcccca     15600
tctggcccac agactggaga gcacatctct cagcaaccac ggagcgccaa ctgcgcacag     15660
ggcctggtcg tcagagcggg gcaaaggcac tgaccgtcac ggccagggcg agggaagacg     15720
ggtgggcagg gaccttgggc agagggggaa gaacctggtg cccaggctgg ccctgccttc     15780
agcagtgaag ctgagtgggg aggcgctgat gcaggggggcc agaaagggct gctggtcagc     15840
cgggaggagc cccccacaga ggaagcagcc agcccagacg cagatggcag ggtcccctca     15900
acaatgtcct ctgaaaagga gaggcgggga ctgctctggt gacacctaca aatagatagt     15960
cagccctcag ccccctgcca tacttctgac aaagcagagg cccccagggg aggcgcaccc     16020
gaaggtacct gcacctgtcc cccagactcc tagagcccac ctgaccccat cccaccaggg     16080
ctccagctac aaaataaatg ccgaggccag ctaggcaagg acgcacactc ggtaccgact     16140
gaataggctc cacgttgtca tgagcgcaac ccacaggcca ccaggccaca ctatgcagag     16200
ctgagatggt ttcggccaag cagcctctca gctgagctga acaagtccag agtccccggg     16260
gggtcgtcac tatggagtaa caattgcgat gcgatggtaa ccctaacagc taaccgtcac     16320
tgagccaggc cctgagctag gtacttttca acgctgcctc tctgcagcct caggacgagc     16380
ctgtgggagc ataaagatca ttccctatca cggatgggga aactgagctc tgaagcagtt     16440
aacgtgcttg tcccagaccg cagagctagg agcaggacac aacagcaggt caggcaggaa     16500
cgggtgaggg gggcctgcat gggcttctct ggaggctgcg catacacgca acccccagga     16560
ccccgacccct gcacctgcag ctcgctactg ccccctcagt gactccagca aacctcgggg     16620
taggggaagg aggctgggaa tacctcgggt gtccgaaaca gcagcttctg cttggaggcc     16680
actgctgcat aatggttgct gcccagcaca ccccaagcca cctgtgccac ctgtggtgac     16740
cttccagcat gccttggtga ccaagctggc cttaggtgct gtgggcagcc aagaatagaa     16800
cagggcccac ccctcctctt cacactaaca caaagcaaga ggcgggcact tcgactgagt     16860
gcatccctct agctcaaggg cctcacggat cacaggggtc agggcaagat cccaattctg     16920
cattcccgtc tgcctttcat cctgctctgc caacaacagc cagtgaggct ggggacatcc     16980
ctgaacctgt ttctcacctg aaacacatca taccattgga ccccagccct ccgggagagg     17040
ccctaatccc tgactgtggt gagatcagat cactggttaa gtacccagaa gggccttggt     17100
caggggctcc aggggtgggg ggtgatgggc gtggtggtat cccgctctgg gctatagtcc     17160
accctgatgg aggaggtctg tggtcagaac cgggctgtgc agggcacagg agcccagagg     17220
gacccccaga gctcacctgg tggtctctga gcagggctcc ctcaaccctc agagaaaagc     17280
acagcaagga ggccgcccag agcccagcgc ctagcaccca gtggcgtgcc agacctgcct     17340
ggatcctgga gatctctcat caccctccaa gtcagtcatg cccaacccag ggacccacag     17400
cccacggggc cgtgaaggtg tgctgagtcc aagaaggcct tcgacactgg gaagccaagt     17460
ggcacctcct ggtgtggagc aggcggaatc ccaccagcct ctgctctgcc agtgggcaca     17520
gctggacgat gagcagaagg ggctgttgct taataaacgt catttcctta agaggataaa     17580
acctttcaaa acagatggaa attttttttt aattaaaact ggtggccaaa gagatggaaa     17640
gcacccccttg tgcctccctc ccatcgtgac ccatcctctg cacacctcaa gctgttcgct     17700
gcccaggtgt ctcctgaggc actgggggcg ggtgagaatc cgtgagccct cggccagccg     17760
tggctctctg gagctctgcc ccaggccatc agggcacacg ccgggcaccc tgggggccac     17820
acagggcaga gcccagctgg gtcagcacac agggccacac tgggcacaca agtctctgag     17880
cctcccctgt ggacgcagct ctcactatcc caccccacta ggtcccgggg atctgtccca     17940
cagggtgata tgctgtcaca gaccactacc agagccatgg cctgctgttc cgcccgcagc     18000
caggtagtca cttgctccac agggacaggc aacgccgcac ttgggggctg ctctgcggca     18060
ggactagagc tccagcagct cagccctcct gagaaggaga actccatgct ctaagaggca     18120
gacgcagcgg acggcaccaa agccaccaca agcccacggg gccctgcatg gcaggtcagg     18180
agtccctgac cactcgctct ttgtaaccag agctgcagtg gagtctacga ggcaaggact     18240
gtgggcggca gtggccacag caaatgaatg agtgtcccaa gggagcaggc ggctgcgggg     18300
aggcacagcc gggacccagg agtcctccgg cactgcagca aactccctgg gccccctgag     18360
cagcgaccag gtggcaagtg catgaactcc cgggggcata acctgggagg gtgacactct     18420
```

cttcgtgttc aaattcttga gaacgcatta aaaatatcac tcagtcacct actctatagt 18480
tttaactcaa aagtaccaaa gtagccaggc gcggtggctc acgcctataa tcccagtact 18540
ttgggaagct gaggcaagag gatcacttaa gcccaggagt tccaaatgaa cctgggcaac 18600
atggagggac cccatttcta caaaaaaagt gttttaaaaa attacctggg cctggtggtg 18660
tgtgcctgta gtcccagcta ctcaggaggc tgaggcggga gaaccacatg aacccagggg 18720
aggtagaggc tgcagtaggc tgtgatggca ccactgcact ccagcctggg taacagagtc 18780
agactctatc tcaaaataaa tttaaaaagc accaagccag gcttggtggc tcacacctgt 18840
aatcccagca ctcagggagg ctgaggcaag tggatcacct gagtcagaag ttcgagacca 18900
gcccagccaa catggtgaaa ctccatctcc actaaaaata caaaaattac ccaggcgtgg 18960
tggcgggtgc ctgtaatccc agctactcag gaagctgagg caggagaact gcttgaaccc 19020
aggaggcaga ggttgcagtg agccaagact gtgctactgc actcaagcct gggagacaga 19080
acgagactcc atctcaaaaa ataaataaat caatcaaaac caccaagact ttttaatata 19140
aacatttatt attccataat tcctttttg catgattaaa aatgtttata taaagtttcc 19200
tgaaaatggt aagaatgcca agtgaaggct gcaaatgccc aagcccccac cgtggcatct 19260
cacggagtct gggccctagg aggctggtgg gtaccacgtg gacccgagac ttcacagtca 19320
agtccctttg gggtacactg ggtttcccac accccagaaa tatgggctct tactgcagga 19380
ccatgggggt cctcacactt ggcccagaag ctgtcacata gccagacagg tgttctacaa 19440
cctaggctag agggagctca tgctccagca gaattcgagc cagaggaggt aaaagatggg 19500
taagatctgc tccctggaca gatgaggcct tggcctcaga acagttactg atcatctacc 19560
agacatcaca ctagaggcag aggggcgcag acgaagacag cccctgtcct caaggccctc 19620
ccaggttggg tggaccatgg aaggttccag acagatctgg caagagaagt gcccacacca 19680
ggggcagaag atgggcaggt ctgctcaggg cggcacggcc tgccaggcca aaaagttcca 19740
acttcagatg ctggagaatg ggcacgactg tctgagaaag ggaaggatgt gatgaaaact 19800
acttggagaa aaattaatct ggccagagca taagataaat gggcaaaggg gaggttccag 19860
aaagcaagga gaccaagtaa aagctgatgt cattggctct gaatctaggc tttcactgaa 19920
tatgcaccgc agggcctgta ggtaaagcct cagagcccag ggagtctgag tggaggagag 19980
ggcaggggac agagctgggg cctgtgtcta cagtgctcag gaggaatagg catggacgtc 20040
agctcggagg ctccagctga agtgaggagg cggccagggc agcacggcca cgcccggatc 20100
cagactcctt ttgggaagca agttcgctct gggggaaagt ttggagaaat ggcctttacc 20160
cgcagaagca agccccagaa catatcttgc tccaaaacta tctcgtacag tgaggacgtt 20220
aagcttcagg tcccctagag gagacagtct gctccttcct ggggcagaac ccaaggtggc 20280
cagagcctgg aaggcaccca gcacccaggc tggtgtgttc cagcccaggc cacacgctca 20340
gatagctatt aatgccccgt tgagcaattt cctgagagct ttgccaggca ggtaccgcct 20400
ccccatctga actaatacag gggtacatcc caaggaagaa atgaaaggtg cccacatttt 20460
gctctgggat taactaggga ggggagtgat aattaactca gtaattatat ttgccatcgg 20520
gctaatgcta aaattagtgt gcattagaat ttctttcctg agcagacacc ggagtgagtt 20580
gggcagcagg agtggctcgg gcaagtcggc acaaagggca cctccagagc cttccacaaa 20640
tgtcagcaaa acccacaaat gtcaaggccg gctccactgc acccagcaga tgaattcact 20700
tccacagcct gagaccgcca gctcatcgga ggccatttaa aatccagccc tctgacacct 20760
gctggatatc accatttacc gtccccagat caagagatca aagggtggaa cctgatagga 20820
cggctctgaa gttcaccaca aaagcataaa cgtgcaagca gagccaatac gtcttttgaa 20880
aaggacaatg aggtgggaat ttacataact gatcttaaaa tatgttctga tgcttcagag 20940
atggagacag cagcattccg gtacacaaag acactcacag gcagtggagc acagtgaagg 21000
gtctggaatc aggacccagg tgtctgtgga cactacacat aaaagagcag catttacaat 21060
gaatggatag gatggaccat cccaccaagg tgttggacaa ctccctattc actggccaga 21120
cccctacctc ataccatata caaaaaaaaa aaaaaaaaaa aaacccagac agaataatgt 21180
ctgaatgtaa aacataaaac agtaacagtc ctggaagaaa ataatggagg atatatttat 21240
aatctggaga tggagtaaca agggatagga aaaaagccat agggaaaaag tagagttatg 21300
attatatgaa gcttcttaat atctttatga taatgtacca ccagaaacaa ggatgaagga 21360
ctagctacag accagcagtg aaacctgaaa caaacagaac aaagaattaa agtccatacc 21420
aaataaagac ctcccacaaa tctataagaa aaagataaac aggctggcac cgtggcttat 21480
gtctgtaatc ccagcacttt gggaggcgga gatgggtagg tcacttgagg tcaggagttc 21540
gagaccagcc tggccaacat ggtgaaaccc tgtctctacc aaaaatacaa aaattagcca 21600
ggcgtggtgg cgcatgcctg tagtcccagc tacttgggag gctgagccag gagaacagct 21660
ggaacccggg aggcagaggt tgcagtgaac caagatggca atcgcgccac tgcactccag 21720
cctggaggac acagcgagac tctgtctcaa aaaaaaaaaa aaaagaagaa gaagaaaaaa 21780
gaaaagaaaa agacaacaga aaaatgggcc aaggataagt gtaggcaatt tgcagaaaag 21840

taaataccaa taaaccagaa atgagggttg tgcaaatcaa aaggtgttat aatttttaac 21900
caaactggac caaagaaaac accaaaaacc aaaatcttgt aattgccagc atcagagagg 21960
atataggaaa gtgtgtgttc tcgtagatgc ttgcaggtat gaactgctac agccttttag 22020
gagttatgta tgtatgtatg cttgtatgta tgtatttgag acagggtctc gctctgttgc 22080
ccaggctaga tctgttgcag tgctgtgatc atggcttact gcagccttga cctcctgagc 22140
tcaatagatt ttcccacctc agcctttcaa gtagctgaga ctacaggagt gtgcaatcat 22200
actcagctaa tttttaaat tttttgtaga catgggggt ctcccaattt tgcccaggct 22260
ggtctcgaac tcctggactc aagtgatcct cctgcctcaa cctcccaaag tgctgggatt 22320
acctggatga gccactgtgc ccggcctcaa tatctttaaa aacagaaatg gacacactct 22380
ttgactagga atgtatccta taaaaacact tatacacatg cagagacaca cgagcaagca 22440
tgctttgtaa tagcaatgaa ggctggaaaa actcctcaat caggtaaatg ctgtcaagtg 22500
cacctgtgta ctatgaaatg gcacttggct tttaacaaga gcaaagacag aaaagcaaaa 22560
gtacaaagta gggtgtgatg gcacatgcct gcagtcccag ctactcagga ggctgaggca 22620
ggaagatcct ttgagcccag gagttggagg ccaggagctg ggcaatagtg agaaaaaata 22680
aaattaaata ataataataa taaaataggc tgggcacagc ggctcatgcc tgtaatccca 22740
acactttggg aggctgaggt gggaggatcg cttgatccca ggagttcaag gccagcctgg 22800
gcagcaaagc aagacaccca tctcaacgac aaattttaaa aaatcagcca ggcaggctgg 22860
gcatggtggc tcacgcctgt aatcccagca ctttgggagg ccgaggcagg cagatcactt 22920
gaggtcagga gttcgagacc agcctggcca acgtggcaaa accctgtctc tactaaaaat 22980
acaaaaatta gctgggcatg gtggcagatg cctgtagtcc cagctactga ggcacaagaa 23040
tcgcttgaac cagggtggca gaagttacag tgagccgaga tcgtgccacc gcactccatc 23100
ctgggcgtga gtgagactcc tgtctcaaaa aaaaaaaaaa aaaaaaaaca aggagccagg 23160
cacggtgggg tgagggaggg cacagaagca gcgcctcttc tgggggcacc cccaatctct 23220
agcgatccag aggcctcagg atcctgaagg gagaaaaaac gtgaagctcc gtgctagaag 23280
agaccataga gattggaatc agctggttct attttacaaa aaaaggaaac tgaggccctc 23340
agaaggtgag tgcctctcaa tgccccacag ggaggcaggg agagggctct gagccctgca 23400
gggccctgga ttcttgcaat ggggtggagt ggagcctgtg ccgcccccac caggcacctt 23460
ctcaggagag gagccgttgt catatccttg aaggggtcct tgagcccctc aaaaggctaa 23520
aaaccacttt cctccttgag tgaaccttca cctcagttta accacaagaa aaactacatt 23580
aaggcccagc gcagtggctc atgtctgtaa tcccagcact ttgggaggct gaggtgggtg 23640
gatcgcttga gcccaggagt tcaagaccag cctgggcaac atagtgaaac cctgtctcta 23700
caaaaaacaa caaaatcagc tgggcgtggt ggtgcacacc tgaggtccca actacttgcg 23760
ggctgaggtg agaggattgc ttcagcccag gaggtagagg ctgcagtaag cggtgactga 23820
atcactgcac tccagcctca gcaacagagc aagactcaaa aaaaaaaaaa aaagcaggcc 23880
gggtgtggtg gctcacgcct gtaatcccag caccttggga ggccgagcgg gaggatcagg 23940
agatggagac catcctggct aacacggtga aaccccgtct ctactaaaaa tgcaaaaaat 24000
tagccgggcg tggtggcggg tgcctgtagt tccagctact caggaggctg aggcaggaga 24060
aaggcgtgac cctgggaggt ggagcttgca gtgagctgag atcacaccgc tgcactccag 24120
cctgggcgac agagcaagac tccatctcaa aaaaaaaaaa attaaatctc aaaaaaaatt 24180
acattaaggc aaactaaaag atgtttaaaa tatatatatt aaattaaata cactccaata 24240
gagcaaatac gaaaataccc agaaaacaca atccccgcac ccccaggaca acctcccagg 24300
gggtccacag caagagaccc caagcacgag agacagagaa cagtgtccct gtggcggaac 24360
ctctggccca tcaggctcta ttagaaaata aggctcttgc cactgagaga aagaggcaca 24420
gtcgcccagc agccacgggc tctggcacac cacgagtcag gccagcaaag tgtcaactgc 24480
cccctacaag gtgacaaact aggacaaact ggaaaccaga ggctggacct ggagcacagg 24540
gaccaccaca tggggctggg gaatgggcag ggacctcaga gcgccaccca catgcctaag 24600
agcagcgcgt atgcgcatgc ctctgcatgg cttagggaca cagggagctc cccccacccc 24660
caacccagga aggcagcccc cactacccag gtagggaacg gataggacca gcaccccgtt 24720
ctgctcgtaa ctcagggctc caggccccct cgggggcaac cagcacagag ctcagacccc 24780
aaatatcttc acccacctcc tggtccccat ctggacaagg gtgctgggga ctggctctca 24840
gtcacaccct cggggtactc ttcaaaggac agctggatgc cccagggcag gagcttttgg 24900
cccccagctc cctcacccca gacaccagct cttgggaccc caccagcatg ggcaaggtgg 24960
acaccatcgt cccgattttg cagatgagga aactgaggct gagggctggc acacggctct 25020
ccagagctga agagaatgca gagagcagcc ggagccagcc ggtgggtccc tgaggccggc 25080
tcgtagcaag ccacagctgc ctccgcccat cacacttgga cctcactggc cccaggacag 25140
ccctccaggg cggcctggca cagagcccac accctgctgc ttcctgaaca aataagtgaa 25200
caaggccacc aagccgagga cctggatgta gccccggctc ccgccagggc ctccccaaca 25260

gactccccat ttggagagcg cattaagtgt ttccaaagcc tcacaaacca cagatgtccg 25320
gctgtctcac ggcttctgta acctgaactt ggccctcact ctgccctccc agcactcctc 25380
tcagggccca ggcccctcct ctgagatgcc agcactgact ccccaacttg tccccatcac 25440
ctggctcgtt cctgaacctc ggcaggagag tctcaggcca gatcctccca ccagccacct 25500
ccaccaggat gcaggaggca tgagacctgc tcgtgccggc tgggagatgc aaccaaccaa 25560
gatcaatcca atcagcggat gaactgacaa atataatgtg gtccctccac acaatggaat 25620
attattcagc cacaaaaagg gctgaaatag gccgggcgtg atggctcaca cctgtaatcc 25680
cagcactttg ggaggccgag gccggcagct cacttgaggt caggagttca agaccagcct 25740
ggccaacatg gtgaaatccc gtctctacta aaaatacaaa aattagctgg gcgtggtggc 25800
gggcacctgt aatgcaagct acttgggagc ctgaggcagg agaatcactt aaacccagga 25860
ggcagaagtt gcagtgagcc aagatcgcac caccgcactc caacctgggc aacagagcaa 25920
gactccattt caaaaaaaaa ataaaaggct gaaacaccca tacgtggtac tacttggatg 25980
actcctgaaa acgttacagt aaccaaggaa gtcagccacg aagacgcatt gtaagattcc 26040
cttcatgcaa aatgcccaga acaggcagaa ccacagaggc agaaagtcga ctggtgttca 26100
ccagggggatc cggggagagg gaacgggaag tcaccgtgta atgggtatgg gttttatttt 26160
ggggtgatgg aaatctctta taacttgata gaagagaggg ttgtaaacac tgtgaatgta 26220
ccaaatgcct gccttctata ctttaatatt ttatattata taagtttcac ctcaatttaa 26280
aaaaaaaaca actcgacacc tttcacctag gaaagatctg gctttagctt gcatttcctg 26340
taactcctgc ctaaagcctt ccagaagctt ccgctgcctt gtggatcaca accagactcc 26400
acaccatgat ctggcctcta agggcctctc gcaggacacc ccgagggtga aggagcaccc 26460
gtgggcccac ctctgcatag ctgcaaagct tctttccctg tcctcccctc tacatgggaa 26520
gctctgcccg caggggcggg gccttatctg ccattctatc gcactcaacc ctagcacttc 26580
actcggtagc agacaccaaa gcaaaacagc aacagcatta taccgggcca ggtgcacgtt 26640
aactcactga attcatggta ggaaggattc tattcccatt ttacaggtga gaaaactgag 26700
gcacacaaag gtagcatcag cttcctaagc ctcccagcac aggaagcggc caggctggaa 26760
tcagaccctg ggcgcagggg ctctgtccac agtgctaact aactactcct gcccccgagg 26820
gctgcagcgg tgagtgagtg agtttgtcag tggactggat gtccaaggtc atacaggaaa 26880
aatccagact attgtaataa cagcctctag accggctggg gccagaaaga tcgaggacgc 26940
tgacacacaa ctgcgctcac tgcagctctg ccagggatgg ggctaaaggt ctcacacagg 27000
gcagttaggg ctccccatag cctgggagag gaacggggtg agataacaga aactaggtat 27060
ggtgcccgaa gtcaaacagc cactgagcat gtaaacccag gtgggtctga ccccaaaccc 27120
ctccaccccc atcagccctg caacccgtcg ctgcaaggga gaaagcaact cagaggcctc 27180
acctgcctac atcccccacc cgtgtgtgtg agttctacta aatgcctgag cagtgacaca 27240
gcacggctga aattaaacgg gttccaaaaa cgacaggaag cacgaagtga atctccccag 27300
gaaagtgctg aacaaatgct ggatcgggtt caccggcgaa tttcttggaa ctgaagaggg 27360
gagctaaaca cacggggccc tgctttggag ggactctct cagggtgctc cacacagcac 27420
ttggttaacc ccactcagcc cttctgggct ctcccagagg gcccggcctt ggccttgggc 27480
atctacagga ggaacctcca gggggagagg gggtgcctgg acaggccggc cctggaacaa 27540
gcacttgggc cccgaggaga gaggactagg gcttgggagc tggggaagtt ctcagcactg 27600
ggaccactag aacaaagcca tttccgtgcg ttcacagctt ccaattgcaa caggaagcaa 27660
tcaggaaaaa taattagcgg cccacttact ggcttcgctg aggtccgagg catgtatttc 27720
acacagtaaa accagggata taacatcaaa accgttctgc agaaagattc ctccctttcc 27780
ttccatttta ggcctggatc accacattca ctggggctcc caggccttgc tgcctaatgt 27840
taaaataatc aactctattt ttgcctcaca cacaactgaa ctctacagct ataattcttt 27900
ctcctcaggg gctcgaacca catggacgac aggcatttga ctccagcaac atcaccccaa 27960
aacgtgcaca aaacccaaaa ctgcaatgag gtgaaaggca acgcggtcgg cctagaaacc 28020
ccccctttaa aacaaacagt ttccccaaaa cccctttttgc ctccttgacc caggcatttc 28080
cggaaaaagg agcggcgctg gcctgtactc cccagatact gtcgctgttt tgtcttcacc 28140
ttgttttgct agctccagac aaggccccac aatgtaaaca cgctcctgaa agaggcagat 28200
ttggggtgaa actgtccata gaatctctag gcttgggtca gaggcaggag gacgtgaaac 28260
aaactccaag ctcctcctgt tccccgctgt cccccacacc tccaagcaga ggctgcagcc 28320
tgggggatct gactacaggg ccacccccgct gcaccattca cactggaaat attcagggag 28380
acagctgttt gccttaagga ggcccagaca aaggggcccg aggtcctccc cgctaaactg 28440
ccacaaacag aacaggagcc gcggcgtgca caggcacttg cggccgtgcc acttggccag 28500
ccatactcca gaaaaacaaa acacgcacat ccgaagagaa tgatttaggt agcaagaggc 28560
ttgcttgaaa aaccacatgg caatctccaa attaaaagaa catgtgtagc gtttcacgac 28620
tgcttaagtt tcctgagtcc tcctgacctc aactccaccc cctgggaaac accaaaagtt 28680

```
ggatgagaaa gttccccccgc cctacctctc cccacgggag tgtacaactg aggcacaagc    28740
ctgcctcccc cactgccccg cgatctggga ccacgtctcc tccgcgtagc cgacccgggg    28800
atggacacta tctggggacc cggcggccac acggggcatt cgggtcgccc gggcacctgg    28860
caggtgtcag tccgcttgga aacccacagc cacgcggctc acaggagcag cgccaccggc    28920
taggccgccc cgcgcccggg ctcagaactt tctcgctgcc acttcagccc gtcctcggag    28980
cacgcggggc ggccgcgcgg ccgctggaaa caggcttgcg aaccggctcc ccgggccagg    29040
cccgcctccg cgccccaagt ccccgctcgg tgcccggccc gggccacacg ggcccagcgc    29100
gggctcggct cggctcccgg cttcccgcgg gctcgggcag gtgaggaccc gcccgcgccg    29160
cacctggcgg agcgggcgcc ctcctcgcca gcccgggacg cagcgtcccc ggggagggcc    29220
cgggtgggga gacaaagggc ccgcgcgtgg cggggacgcc ggggacggca gggggatccc    29280
gggcgcgcgc cccaactcgc tcccaactcg ccaagtcgct tccgagacgg cggcggcgcc    29340
cgcgcacttg gccgcggggc cgcccgggcc attgtccgag caacccgcgg cccgtcttac    29400
acgccgggcg cgggaaggta tcgaatcagg                                      29430
```

<210> 8
<211> 33769
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> (33739),(33749),(33758)
<223> Identity of nucleotide sequences at the above locations are unknown.

<400> 8

```
cttcccctta cactggtcct tcgacccgcc tcggatgaaa actgaatggg tttagcctta    60
gaggctctcg gtctctaagg gaggtgggtc aggatgccgg ggacagggtc ctcttcctgg    120
ggcaacgtgg gggaacgagc cacctacccc tccactgaat tgccctgggg tgtgggtacc    180
gacggctcat tcggtgtcca gggtctgaga tgtgttgaca ggaagaatga aaggggatgg    240
gagggatggg gcgaaagaag ccacctgcag ccccaggaac tatctggcca gcacaccgtc    300
acccagcggc ctgagccacc cctgccagag ccaggaggag accctgccaa tgggtcacca    360
gtgtgcagga actcagaagg tcatcacagt taataccctc catgccccaa tgtgggaaaa    420
caggtttttt cacaacaaac aagataattt ttgttatttt ggcaaaagga ggcagggcag    480
ccccggacac ctccatccca cctcatcacc cagccgcagg gccccggcca tccctgcaga    540
cagagtggat gtcacaacct ccctgcaccg aaccaagtgc agctcccagg ccacaggcca    600
cccaggaaag gtccagtggc ccccggaggc tcccaccgca ggcctcccac cacagccggc    660
accaacccag gatagctgtg ttctcctggc ttcttttcac acgggtagca gaaagctgag    720
atccggggaa agctgagatc cagggaaagc tgagaatcgg cctctgctgc ccggacgccc    780
accccccagct ctgctcccag ctccagggcc tccttctcag gtgcccttac aggaggcaga    840
gggcttgagc cacctcctgg gcctggggca cgcaggatga acggggtcac ggtgcaggcc    900
actgtccact gcgcagatcc caaggccata aacagcctgg ccacagtggc ttcccagctg    960
gcaggcggcc agattatttt tgttgtttag caattgatta agtttctccg ctgcccccag    1020
gggtaagtgg tggggcaaat gccgcaaccg cagcatttga cccgggatcc tgtgccaagt    1080
gaccataggg tcacaaagca caagggaagt ggctgggccc gatgctggct ctgctggaac    1140
ctgaggccgg ccactgtcac ctgcacggtg cctgggacct tccagcaagc acagagaagc    1200
tatggccctc caggagcagc tggcaggcac cttggcctgc agtcaggggc tctgtctgct    1260
cagctctaaa acaggaaagt cgctgctctg cctggggtca gggcagccag agagtgacca    1320
agtcagtgcc ggcctcagga agggacctgc aggcgggtcc cttcctctcc catccctcgg    1380
tgccagccag cccctcctgt ggcccccccac tgcctgcctc tgcccccatg ccccaccaca    1440
acctcaggcc catggctgca tggccactcc ccaggcaggc agtggggatg ggatttcacc    1500
atgttggcca ggctggtctc gaactcctga cctcaggtga ggagttccta aagtgctggg    1560
attacaggcg tgagccaccg cgccagccct ccctgtggta ctaaacactc acaccccctt    1620
gctggggacc ctggtgaggg aacacagcct cacaagtgaa gtgtggtttt gttgagcaaa    1680
tgacgcctgg gcagccctct catctttgcc taaaactgaa gaatttaggg gcgtggatgt    1740
```

```
ataaaacagt tggtgactta aatgaaaaag aaggccacac tcccccccttt aggcaggcgg 1800
cctaattctt taaaagccag cacagggtgc ctttctgaac ccaggcacac agtaggtgtt 1860
caatggacag cagcggttac ttgtactgct catgacaccc tgtctgtggc ctctgcagct 1920
ggctccagcc tgacgcatgg ctgcgcccct ccgcaaggcc accccggtat acatggaaac 1980
tctgtggaga aggccttggg ggccggccag gacgccaggc ccagatccca tctgcgccct 2040
tcctccatag acctcagcga gctctcggca ccatgtgcct caggcccatt taagaagtag 2100
ggccggccag gcatggtggc tcatgcctgt aatcccagca ctttgggagg cccaaggtgg 2160
gtggatcacg agatggtcag gagatcgaga ccatcctggc taacacggtg aaaccccatc 2220
tctactaaaa atacaaaaaa taagccgagt gtggtggcgg gtgcctatag tccaagctac 2280
tcgggaggct gaggcaggat aatcgcttga gctcagcagg cagaggttgc agttagcgga 2340
gatcgcgcca ttgcactcca gcctaggtga cagagagaga ctctgtctca attaaaaaaa 2400
aaaaaaataa aaaaaagaag cagggccagc cacggacgac ccctcacaca gctcccagga 2460
cgcgtgcctg ggtatagggc tcaggaccat gaccgctgca gtggccccca agaaacgtta 2520
cttttgtcac ccaccccgcc tcagtggcag tagccaaaat aacggattag aatggaacca 2580
tgtgacaatg ccactgcccc aactgacaga agatggctat cagcagttca cgcggcccca 2640
cctatcacaa gtgcagggca ctctacaact tatgcatcct tccccagaca ccgtcctttc 2700
gaccctccca ggtcagcaag gcacacaggg cctacatttc acagccacac agcagagggc 2760
tgaggctgga actcggatgc tctgatttcc gttcaatcac atccccagag gtggcacaga 2820
gacggggggc ttctcttgac aaagtcaaga aagtcactgc cagctccact gaagaccaaa 2880
gaacctcagc tctcaaaccc tcttgaaggt gttaccgaac tctcccagcc tgtttcctgg 2940
gtcccgatgt tggtcccgtg ggacacagga agaggaagaa gctccctaga gcagagcctg 3000
gtgcacctgc cacactctca gagggctgcg cacgggcgga ggagccgtgt gcaggagtgg 3060
ggtctggatg gaggggcgct gtggccgggg gcagggggca ggggaagggt gctccaggtg 3120
gtgggcacag cacgagcagg ggcagggagg tccacactca gatgtgcaca gggagaaaca 3180
aatcgtgcat ttccattgga ataggcggta aaaggtagaa aaacagagtg gggggccagga 3240
agggagtcgg agccttctag tgtctctctg caggtgagcg gcagcccgag gtgtcagctc 3300
agcagacttg ggtccagggg gccgtgtctt ctatcactga ccccagggca cacggaactg 3360
gggagggaga gcagaggcac agggcacggt cagtgaaacg aaacaaggag tcatcaccaa 3420
atgcggaaag ggcaaggagt gcccgcagcc gcacaagggt tctgtctggg caacgtgggc 3480
gtcccaccag gccccgcacc ctgcaagcgc aaagctcgcc actgaagata aagggaagct 3540
gttggagctg cggagctggt ctggggtccg catggagctg ggcttatgct gcagtcacaa 3600
gggggacatg gaagaggctg caggggacaa aaccagtgac cacagtctaa ctctgagcct 3660
gtggaaaggc gcccacagca ttcacccatc ccagagatgc cattccccct gtgcccccgc 3720
tccacggtga cagcgttctc caggaatatg atgcgcccct ctcctcttgc atcagccctg 3780
acagtgagta ttcaggccaa aaagcagaag agcacagctg cgtggttcca tttccatgta 3840
gttctggaac aggcaacgct aatccaaggt gatagaagtc aggagagtgg tggagggggc 3900
gggggttgag gatggcaaag gggcaccggg aactttccca gtggtagaaa tgttctctgt 3960
ctggaccgtg tggtagttat gcagacatat gcagctgtca aagttaatcc aaatgtacac 4020
gttaaaatgt gtgcgtttta ttgcctgcaa gttatacctc aattaaaaaa ataaagttag 4080
cactcaggct tcttccacaa cttcctgaac cgtgtgagct gattttcttg ctattaaaaa 4140
ttcacggtcc atggctgaga acagcagctg ccttctgttt gcaaagtcaa cgccaatcac 4200
tgcccggccg cggcagactc ggccccacag gacctccttt ctttttcccc tttgacctac 4260
ttccctgata agtgacaaga cagccagact ctgggaacaa acgcccgtta ttcggccccg 4320
agctgagcgg gccctgcttc ctgagctaat ccgcccggac agacggaggg acgtgagggg 4380
ctttgccgtc ggctccagct gtcagtctgc ccgtcagact cgacagtggc cccctctgtt 4440
cctcccgctg cccccactcc atccccgact tctttttgtt tcctgtccct gacagacgaa 4500
catctgttaa aactctgtct gggtgagctg tggccagcgg cccacaaatc cccaagccgc 4560
accccagcct catctgggcg ctgccgggag cactgcctgg ccaccctctg gacatagctc 4620
tgagagccac cggccagggc acgtgtggcc cgagtggcat ggtgcacgcc gctaagccca 4680
ctgcccaaag gcccccaagc aggagggatg tgcaggagac aaaagtcaaa agaacagggg 4740
cacgttccac agaggatggg gctggagggg tggcagtgag gaacagcagc ttccgaggat 4800
ggcggtggca actcccaaat aaggcctcac tcctgctgtt tttagctcat tccacataat 4860
tggaaaaaca tggcagaaac cgaagccagc tgctgccttg gtcctggggc tgtgtggagg 4920
gggtggggag gccggaggcc caggctctgc actcgactgc tggggatgag agtgactctg 4980
agctgcagag agcagcatcg cagccgccat ggtcccattg agccccggcc acgctgggcg 5040
gcagaggctc gtgggatata cctgccctgt ctcatggggg tcacttcagg aggggcgggg 5100
gagccaggac acagcccagg gctagcggtc accctgcagc tcaggggcca cgtaaatagt 5160
```

gccaccttga aggcacacag cagtgcgggg cccccccgc caccaacgca tccctacctc 5220
taggaggccg cctgtgtgcc cctgggaacg ctgctccctg tcccttgggg tcctggtgtg 5280
accaccctct cagccccttc cttggggaag gcacctgact ccctacaccc agctggcttt 5340
catttgctca aaatcaggaa aaagcagaat tcaagacatc acagaaatgt cttcgcctgt 5400
aactccatga aagataaacg gtcagacacc caggagggag tcccagggac ccttgagtct 5460
cacctgaggc tctggcttca aacctcgaga tgtttccagc catgctagcg ccgcccccca 5520
caacctgccc cacacagtcc tcccttggga actcacagat ttggccccca cctgccccgt 5580
ttcttctggt ggagtgggtg cgttgggttg gggtggggct ggggactctg gatgtgtctt 5640
aagagtctga gtgattctga cacagccagg ccctgccccc ctcctgacct tcgccccaca 5700
ggaaagggag ccacacgcct gaagcgccca gcacaccccc ctccgtcctc cccaggtcac 5760
ccgctggccg tgtgagccgt gctccccact gccccttcac ccaccccagc tcctcctggc 5820
agcacccagc cttggaagct acttctgatt acaaccgccg aaggaagact cgctccctcg 5880
gcactgaccc agacagcctg caccatcacg ctgctcagca caacccacac agccttcctc 5940
caaaccccat ggagcgggga gtataatcac ccccttcta ccaacggaca aactgaagca 6000
cagagaggtt aagtcacttt cctaagctcc caacacgatg acaaaaaata gaaggtcagc 6060
ccgcaagtgg aactaggtgc tccaagtccc cggtctgcct gacactgcac ctcctcgccg 6120
ccacggtccc gggtccgcct gacactgcac ctcctcgccg ccacggtccc gggtccgcct 6180
gacactgcac ctcctcgccg ccacggtccc gggtccgcct gacactgcac ctcctcgccg 6240
ccacggtccc gggtccgcct gacactgcac ctcctcgccg ccacggtccc gggtccgcct 6300
gacactgcac ctcctcgccg ccacggtccc gggtccgcct gacactgcac ctcctcgccg 6360
ccacggtccc gggtctgcct gacactgcac ctcctcaaca ccaccacggt cccgggtctg 6420
cctgacactg cacctcctca ccaccaccac agtcccgggt ctgcctgaca ctgcatttcc 6480
tcatcaccac agtcccgggt ctgcctgaca ctgcatttcc tcatcaccac ggtcccgggt 6540
ctgcctgaca ctgcacctcc tcaccgccac ggtcccgggt ctgcctgaca ctgcactttc 6600
tcaacaccac tccttggccg gctcccaact acaaaccaag ccatgtcttc catcctgaat 6660
cctcttggcc taaacatcac tcacaatgcc tccctcggga acaggcacaa gtcccaccag 6720
cacagcctcc ttcgttacct gcgtttccgc tagcccaggg ccagctccag agccctcacc 6780
acagagcctc tatccttcac ccccggacac tggacctcac caacccatag cctggaggag 6840
atccctgtgt gaccccaggg cctcctctgc ccgactctga atttcactgc ccaacgtgac 6900
acctcggaag gctctctggg cactggcagc cctccatggg caccgctcct tctggccagc 6960
tctgacatcc cggctggtga ggtgccctgc acgaggcctc tgcccactgg gacctcacag 7020
ccgtgctgtc agctgcaaca agcgacagaa tttcacgttt tcttcacgtt gcccctgggt 7080
gagcagctcc aggtagtttt cagtcgaggc gaggcgtccc gtcagcagcc aggcggcaca 7140
gctaattcat gcccgccggg cgcacggccg caataccaat gggcacctgc agcctggaaa 7200
gccacagagg aaccgagaac agcgactgtg ctcaggtgac aggactgtgg tcttttaaca 7260
aaacattttc ctttaacgtg atattttacg gcaaggaatg aaacctggag ggcaggacat 7320
ttggatacta aagccccagg ctgccgcgtg gtctgctttg tgaagtctga agcccgcgcc 7380
ccattctggc cccgctcaca ggtccggctc tgactcacca gcttcaatgc taggccgtgc 7440
ctgtcctcca accagaacat gacttcctta aggacaaagc cgtttctcgc ccatccccat 7500
ctccctctgg attaagaaat atgggaagat cttctagaac cacctcaaat ttgcagagag 7560
ccatcctggt gacaaaccct tgaaatgctt ctaagaagag tttaggtttc ttctcaactc 7620
taaaacctct agaaaactct atttccacac cagctgcccc tggaacactt cagcttcaaa 7680
agggcccagg gcagggagac ggaggagcca gcatccacac cgagcaccag cctgttaatt 7740
aacgggaagc gggtggggcc catctccagg cagctctgag gtcagactgg ggaaccatgc 7800
ttacaaaaaa aagtgaactg aaacgctcac gtcctcatgc aaaaccagac tcccagttgc 7860
atctttctgt ctcattgagg agctttttcc tccctttgac agaacaccct acacacggca 7920
tctggaacca aagcagaaag attcaggctc agagtaaaac agtccccaca ctggctgcat 7980
gtggacgttc ccggcccaga gtctcgccca agcagggcct ataaatgaca caaaatgttt 8040
ttctcctgcg tgccagtcat gctccaactg agttatgtgt aaaagtgcct ctcacggctg 8100
agggcaaaaa cagttcccac aagactagag aaaggtgacc cctgacggct gagtctctag 8160
ggagcgtgga gctgcgtgct cagccctgcg gccctgacgg ctctggaatg gaaaagctat 8220
ccaactggaa gggcagggct cgctgctagt ccagcggtcc aaccccacag gtgtctgtgg 8280
tgtcagctcc atgccacaga gcccagggct ggggccagag ccaccaggcc ccctgccagc 8340
ctgcaggggc ctcctcctct gggtagccta accacccct gtgagcgcag gcagcctcct 8400
ctaatcacca cagggcctgt cccccctct cccccgcttg caggaaaatg agccctgagg 8460
actccccagg gctgctctgg gcctggacat ggagactggg aattacattt gcagaaggag 8520
cgcaatgccc ttgaagggct cagccacgag cagccagtcc ccagggctca gaaggcccag 8580

```
ctgttagaac cctgggagcc agcaaagagc caggggctcc acctaagtct atagcccctg    8640
cctcttctgg ttgggaaaga aatcaacgcc cctttactgg ctcccactga cagcccactc    8700
ccccaggtat gggaggattc tgggacgatg caggcaaacc tggaccctga gtgaacctgc    8760
cccagctctc acgggcctgg caccagccac agcacctaag gcgccggtca tggtgacaac    8820
atgaaggtga taagggcatg gacagtggac atggcagctg gacactgggc acccactgga    8880
tgccaggcac ccagcacggc tccgtcaccc ctggatgagc agtggccctt tgcaagccag    8940
ggtagcctgg gcaagttatt tgggggtctc caagcttgtc cagctgtgcg acttcactga    9000
gccatgagtc tgggatttta tcagggccca cacccgttcc tggaactctg atacgtgagg    9060
gagccacaca gggaccctta acaaaagctc ccagggcaac atgttctctt gcctcagtct    9120
cccaaatagc tgggattaca ggcgcacgac taccgcccgg ctaatttttg tatttttagt    9180
agagacaggg tttcaccatg ttggccaggc tggtcttgaa cccctgacct caaatgatcc    9240
ttccactgtt agggcaaggc acctgacagg cacgactgca cgatctgctt gttgggggct    9300
gtgtccattc cccactcctt cgacaaatgt ccacacccag ccttgctttg acaccccaag    9360
aacagagatg gtgacacctg cttcctacat gcccattgct ctcccaaggc agacatcccc    9420
agcagatgca acacagtgtt taggcagaca tcaccaatcg atggtggcaa cagacaccag    9480
gccctgctcc ctctaactcc agtggccagg ccccaagcca gctctcacct gcccactccc    9540
aacccacagc agcaagactc agaaatggca aaaacacaaa gagaacagaa acgccccata    9600
gcgggaggat gactaaaaga catgtcttga taagatattg ttcaggcata ggccaggcac    9660
agtggctcat gcctgtgatc ctagaacttt aggaggctga ggtaggtgga tcacctgagg    9720
ttaggagttc aagaccagcc tagccaacat ggtgaaaccc catctctact aaacatacaa    9780
aaattagcca gacatagtag cgggcgcctg taatcccagc tgcttgggag gctgaggcag    9840
gagaattgct tgaacctggg aggtggaagc tgctgtgagc cactgtactc caacctggac    9900
aacagagcaa gactctgtct caaaaaaaaa aaaaaaaaaa gatatccttc actaaaactc    9960
atgtctttga tacatattta cctcctgcaa tcgcaaatgc ttctgcagtg cataaagtga    10020
aataaatagc aggaagcctt acggttcgat cacccacaca gacacacagt cacatacagg    10080
aaaaacgcag ggagggctgg ggaacaaaaa aacagaagat aaaatgtgga gacagacaca    10140
ccaagagagt aagagaccac ctccagacct cccttcagct tctcaaacac acgagccggg    10200
cccgttacag aatttgcggg gaccgctgca aaatggaagt gcagacagcc ccttactcaa    10260
aaggtaggaa tttcaggtca acaacagagc tcacctcata tgactacaca ggtcacacag    10320
cccgtgaagt cggtcccaac accagcatgc tcctgcctca aagccgctgc acgtgctgtt    10380
ccttctcgcc tttccctctt ttagtccttc agatctcagg cctcctgaga gagacctctg    10440
acctgccggc tcaggcggcc acacccccag tacaggagtc tccggctcag cccctgctgt    10500
gttccgtacc cgatccaggt ctgtcctatg tccatctgtg tgccggcttg cttcctgaca    10560
tggcccccac cacacgtgtg cctcggggca ggggaacagg cccgtctcat taactgcttt    10620
cttctcagat attttctgga atatttgtgg atattgggca acatatatgc tccacctttt    10680
tcagactagc caggacgagc tgcatttttt tttttttttt tttgagacag ggtctcactc    10740
tgttgcccag gctggagtat agcggcatga tcttggctca gtgcaacctc cgcctcctag    10800
gctcaagcaa ttctcctgcc tcagtctccc aagtagctgg gattacaggc ccgtgccact    10860
actgcccagc taatttttat atttttagta gagatggagt ttcaccatgt tggccaggct    10920
ggtcttgaac tcctgacctc aaatgatcca cctgccttgg actcccaaat tgttgggatt    10980
acaggcgtga gccactgcgc ccggcccgag ctgcctgttt tacacctttg ccatattccg    11040
gtgattctct ctcccctccg tcccccggcc ctgactgtgg tggccactcc ctgccgtcat    11100
gagcccgtat gtcctcactc tttccctttc cgccaggact tcaaccaaca ctgcagagcg    11160
cagggtccag ctccagcact gagttcagcc tcttctcacc aacagacagg caggaaagaa    11220
aacaaactct gagaaggcca aggttcccgg gcagccagca agccaagcat ccttctccgc    11280
tgaggcttgt gcagccgagg cacccccctcc tccagggagc aggcagcgtc ctggggcagt    11340
ctgcgaggga gaccagggcc cttgctccac cagggcccca ggtatggggg cagcagcaaa    11400
ctcatggctc tgggagccag accccacctg ctagaaccta ctatgccacc tgctgtgggc    11460
aaccccaggc tggtgacttg ccctggcctc ctctgtaaac aaagggctca tccaacctgg    11520
tcaaaccact cctccccttc aagggtctat aatcctccct taacctgctt ggtccaaacc    11580
cctggtgtcg ccaggtcact caggaggcag ctcatctgga ctccttccct gggtccagtt    11640
tctctctcaa cattgccttt gaggccgagg tgaacggtca acagcgaagg gccccagagg    11700
tgatggagga gcgggtgtcc aagacactca ccctttctaa tgcactgact ccctcgtgga    11760
ctcacttgtg ccgtctcccc cacccaccca gccccagagc ccagagtgcg agcgccagag    11820
gcccgggatt ctgtctgcac cgcggggtcc ccagtgcctc ggagcaatgc cagcacccgg    11880
caagtgttcg acaaatgcct gctgaatgag caaatggatg gatgaacgaa tgaatgagca    11940
agcagatgaa tgaatggggt gctgtccaga gccgtgagga ctaggccgcc caagtcccca    12000
```

tttctcaaat tctccttctc ccgacttggg aaacaagatg cttggtcggg gaggctctcc 12060
aaccatcccc tgcagcagcc ggcacagcgg acagaccctt tgatgtaaca gccatgtctt 12120
cattaaagat gccctgctct cagaaagaga aagacaaata caaacctgga aaatcctcac 12180
caaacgcagg acccctgcca gggagcagag aaaagaccca cacgccacgg gcgccacgac 12240
cacacacaca ccccagccgc tgcacacaaa cacagaccct agccagcaag aacaggggga 12300
ccaggaaact gttcctaaag tcaggacccc catgtgctca gacagcagtg agagcaagga 12360
cacttctcca tccaccggat gccaggagag tccttttagg gggccccaca ccgagactct 12420
gcccttagga ctgttcctga gtgtggaagc cagcccactt ggaagccccc tgccctcccg 12480
agtgggacac cggcacagga agcaggccct gtcccccacc actttctgca agctgggccc 12540
catcacgcta cagaaacggg gaggactggt cccagggatg gcgctttcct gacacctctc 12600
gttacccccct cgcttgccag gccccagggt cagccccaga ggccagactg gctatcccag 12660
gcccgggagc atccccgaag gcgagctgca tcctgaacgt gtgtgatttc ccgaagggcc 12720
cgccccgaac cgacacctgg aaagaaagat cctcagccgg tgccccagag gagaagagcc 12780
atgcctcact gcaacacagt cccaggaagc accaagtgcc tgaggaccaa ggcggagagt 12840
aaaaaagtgg aaaatatctg gggcaaaaat aaaacaaaac aaaacaggat tgacctcctg 12900
ggctcaagca atcctcccaa ctcagcttcc cgagtagctg ggaccacaga cttgaatcac 12960
cacacccgcc aagtggatca tttcgaacgg gtttgccgag gttccttctg ggcacccccc 13020
ggcggccgca acccattccc gccaggcccc gccccgcccg cccgccccgt cccgtcccac 13080
cgcctcacct gccttacacg tcctgccgtt gtcctgcagc tgcacacccg tggggcaggc 13140
gcatgtgtag aaaggctcgc ttggggacag caggcacagg tgggagcagc cgccattgtc 13200
ctcctcacag cgagtgtgga ctgagaaaac caggacagac tgagagaagg ttccagaaga 13260
ggaccgtcac ttgtttctga atgagtcaca tcctgcctcg tccccgtga cagcctccag 13320
tgtgtccctc tgcccaaaca tcggcctcaa gtggcatcag ggacctcccc gcgggcacca 13380
ttccacctgc ctcatcgctg gccccgtcca catgggcccc tcagcctggc cagacggcct 13440
gcaatttccc caaaaccagc cgtgaccttc ctggccaccc tcacacccag atgtgacctg 13500
cccatggagt gacatcctcc ccatctgctt cctcccacca agctcctatg actagaacac 13560
cctccccagc tcctcggagc cccaaaagga cacccctctg caaaggctgc cccccacgct 13620
ccaatggccg gggtcaggac ctgcctgtgt ggtagtgacg ggaaccccag agacaatggg 13680
ctcctgggca aaaggcttgt cttgtctttg tgctatgtgt ggacccagca gcttccatag 13740
gaacactgtc cttcttgctg ggatggccaa gcttgtcact ctcccaagcc ctcctatgac 13800
caacagcaat tgaacggaac tcgataaatg cttccagcac ctcattcaaa ccaggggaaa 13860
gctgggtgta gcagccccaa aatacggata taactggaac aacaaactca tcaaaatgaa 13920
cctctccctc cctcatgctg ccccaagtgt agatgggttt tgtgaccacg actttctcac 13980
caggaaacag ctccagagag ccccaccctc ctgtgtcctg ctctgggaac agctggcacc 14040
cctaggcccc acatttcaat tcaaagtcca aaccttccat aatggcctgg ccagaaatct 14100
ccatccctgg tccctgtggg agtgggccac tgtccccaga gccgcagccc cactgtcaca 14160
gaagctggtg catttcccca tcagggacct ctgtcacaac ccagcgtggc cccccaggctg 14220
agaactgctg attctgggca gattattcat tgataaatac gcgacttgca gggccaagca 14280
tggtggctca tacctgtgac cccagcactt tgggaagtca aggtgtgagg atcactggag 14340
cccacgagtt tgagacaagc ctgggcaacg tggcaaaatc tctcatctct attaaaaata 14400
catacacaca cacacacaca cacacacaca cacatatata tgtatatata aataaccata 14460
tatatatata cacacatacg tgtatgtgta tataaataca tatacacaca cacacagaca 14520
acttcttctg ggccttgaaa acgaggcaac cttccttgga aatccccttg ccactgctga 14580
gcctgaaata gcccccatga gctctgcaga ggggtcctct gcaggcccgt gtcccccagc 14640
cagccacaca cctccctcca ttgcagcagg taccccttta gagaggggc cccccagagc 14700
atgggcttct gcagggaggg gtcacctgcc cccccacccc acccacgccc gcgcaccccc 14760
acgcccccgc atcctcccac tcccctgccc cgcgcccccg ctccccccag ccccctcacc 14820
ctctccccg tgccccaacc ggcactcaca aaaaggctgc cgctcctggc tcagcacctg 14880
gatgtccatg ggtgagtata gggcactcag gatctccttc ctcttccccc cagtgcgctt 14940
gttgcaggca tggatggagc gggtctgcca gtctgtccag tacagagtgt ccccggagag 15000
cgtcagggcg aaggggtgcg tcaggctgcc ctccaccacc ttctgcctgc agtcagggaa 15060
gcggggtgga ggagccatca ggagggtccc ccgacagtca ttgctgctga cccaattaat 15120
ttcttttttt ttttttgaga tggagtctcg gtctgtcgcc caggctggag tgcagtgatg 15180
taatctcagc tcactgcaac ctccgcctcc cgggttcaag caattatcct gcctcagcct 15240
cccgagtagc tgggatcact gatgcccacc actacgccca gatgattttt gtatttttag 15300
tagagacagg gtttcatcat gttggcaagg ctggtctcga actcctgacc tcaggtgatc 15360
cacccacctc agcctctcaa agcgctggga ttacaggcgt gcgccaccat gccaggcttc 15420

```
ccatttgctt tcaaccagac aagtgaggcc aggtcaagag ccccaggagc tggcgccctc     15480
gtacatttct cccggcgtgc acagggcacc tcccaaacac agcctgtgat ggtgacacac     15540
gggctccccc aggtcaagtg gcaaagtctc ccccagggaa gaaaggagga agccatgcct     15600
ggcaaaaagc acacctctcc tgcccaacgc tttaacctct gtatacaaat caggccatgt     15660
gcactcgctc cttcttacaa tgctcataat ttatactttc agagtaaatg aaacttggca     15720
tcaacccgag aaacagctat tcttttctag atgcttacag tgcccagcaa atgaggactc     15780
gggtgtaatg agattatgga cactggaaac aggatcataa tgtgacgtgg tcggtaatgt     15840
gcagttttat ttgcttaatg accctcgccc cgtgacaggc tccctgaggg tgggcctggg     15900
ggcagaggtc cccgccacgt ccccagccct cagcacagtt gccaggagag ggtgacactc     15960
atgaagtggc acagggaaga tgggagctgt gggctctgca gatccaccac ctcttctgtt     16020
catttttgtt gatgctgttt tttaagaaaa ttattgaagt aaaattcaca ggacatacgt     16080
ttactttttt tttttttttt ggagatgggg tctcactctg tcacccaggt tggagtgcag     16140
tggtgtgatc tcagctcact gcaacctctg cctcccaggt tcaagcgatt ctcccacctc     16200
cgcctccaga gtagctggga ccacaggcgt gcaccaccac acccagctaa tttttggggg     16260
gtatcttttt ggtagagaca gggtttcgcc atgttgccca aggctggtct tgaagccctg     16320
agctcaggcg atccacccgc cttggcctct caaagtgctg ggattacagg cataagccac     16380
tgcacccagc ctaaatttac cactttaaag tgaatagtgt tacctagtgc attcgcaagg     16440
cggtgcagcc tccacttctg tctagttcca aagcacttcc attgccccac aggcaaaccc     16500
cacacccggc agcagtcatg ccccagtccc cgcccccagc cccggcaaac acttttgatg     16560
gacttaacta cacacattct caacatctca tataaacgga atcacaatat acagcctctg     16620
atgtctgtct tctttgactt ggcaccatgt tttcgaggtt catccaggct gtagcatgtc     16680
agtgcttcat cccgttttag gggtgaacca tattccagtg tgcagacaga aaccaatctg     16740
tgcatccatt cacccactgg gggacctttg tgtcatttcc accctcggct gttgtgcaca     16800
gtgctgctac ggacattact gtccattcac attttgtgtg aagacctgtt ttcgattctt     16860
aagagtatac agctaggagc ggaattgctg ggtcatacgt aaatcaatgt ttacgtctca     16920
aggaatcaac aaactgtttt ccacaatgtt gtcttttttg tttgttttct gagacagggt     16980
cttgctctgt cacccaggct ggagtgcggt ggtgtgatca tggctcactg cagcctcaat     17040
ctcctaagct caatccatcc tcctgcctca gcctcctgag tagctgggaa cacaggtatg     17100
taccaccatg gccagctaat ttctaattt tattttttttt tgtttttgtt tttttgagac     17160
agagtctcgc tctgtcgccc aggctggagt gcagtggtgc catctcagct cactgcaagc     17220
tctgcctccc gggttcacac cattctcctg cctcagcctc ccgagtggct gggactatag     17280
tcaccggcca ccacgcctgg ctaatttttt tgtattttta gtagagatgg ggtttcaccg     17340
tgttacccag gatggtctcg atctcctaac ttcatgatcc acctgccttg gcctcccaaa     17400
gttctgggat tacaggcgtg agccaccacg cccgacctta cttttaattt tttaatttta     17460
ttattttatt ttattttttt tttttttgag acagagtctc gctctgtagc ccaggctgga     17520
gtgcagtggc gggatctcag ctcactgcaa gctccacctc ccaggttcac gccattctcc     17580
tgcctcagcc tcccgagtag ctgggactac aggtgcccac cacgatgccc ggctaatttt     17640
ttgtattttt agtagagaca gggtttcact gtgttagcca ggatgatctc aatctcctga     17700
cctcgtgatc cgcccgtctc agcctcccaa agtgctggga ttacaggcgt gagccaccgc     17760
gcccagcctt tttttttttt tttttttttt ttttgagata gagtcttgct ctgtcgccca     17820
ggctggagtg cagtggcggg atctcagctc actgcaagct ccgcctccca ggttcacgcc     17880
attctcctgc ctcagcctcc cgagtagctg ggactacagg cacccaccac cacacctggc     17940
taatgttttg tatttttagt agagacgagg tttcaccgtg ttagccagga tggtctcgat     18000
ctcctgacct cgtaatccgc ccgcctcggc ctcccaaagt gctgggatta cacgcgtaag     18060
ccatggcgcc cagcccatgt ggccattttt cagtgagaga agccagaggc ccatcactct     18120
cggttgctcc ctgggccatg ctctgcctca gccagaagca ctgagggaag gtcagcctcg     18180
gcccttgccc cagccacagt cacagataaa ggggcctgca caggtctgtg tggctccaga     18240
gctcgtcacc caacacacga cgcttccatg tgaatagccc caggtgcatc atgaagagcg     18300
atggccgctg cagaggcaga agaatcccgc ggggaagcag gtgggagaga ggctgagaac     18360
agaccagacc ctggagctac agaccctatg ttccaaccct ggctgggact agctgtgtgg     18420
ctctgggcaa attcacatgc ttctctgtgc acaggggatc aaaatagcaa acacaggcta     18480
ggcacagtgg ttcacaccta taatcccagt gctttgagag gccgaggtgg acacatggct     18540
taagctcagg agtttgagac cagcctgggc aacatggtga aacctcgtct ctacaaaaaa     18600
aataccaaat aaattagcca ggcgtggtgg tacgtgcctg tggtctcagc tacttggaag     18660
gctgaggcgg gaggaacact tgagcccaag aagtcaaggc tgtggccgcg tgtggtggct     18720
cacgcctgta atcccagcac tttgagaggc tcaggtgggt ggatcacttg tgatcaggag     18780
ttcaagacca gcctggccaa catggtgaaa ccccgtccct actaaaaaaa tacaacaatt     18840
```

tgccaggcgt ggtggcgggc acctgtaatc ccagctactt gggaggctga ggcaggagaa 18900
tagttagaac ttgggaggtg gaggttgtag ttagccaaga tggtgccgct gcactccagc 18960
caggggggaca gagcaagact ccatcccaaa aaaaaaaaaa acaaacaaac aaacaaaaaa 19020
agaggtcaag gctgcagtga accatgattg tgccaatgca ctccagcctg ggtgacaaag 19080
tgagaccctg cctcaaaaca ataaaaatat aaataaaaat aaaacataat agcaaacgtt 19140
tcatagaggt ggtatgagca ttaaatgaac tgataaacgt ccctggaaaa cagtaagtgc 19200
tatggaagga ttcgctgccg ccaccgccac caccattagc atgtttcaac ctccatcacc 19260
ctcactgtcc cctgtcacca tcctttgacc agggcactcc cagctgcagc ctttctatcc 19320
tcttgtccac ccttcataac tgtaagatca ctcagctccc aagaaccaca gtctacaggg 19380
taaccacatt tccaaatctc aaaccagacc cgctggtctg cacttccagg gacaacagga 19440
tattttcaaa ccagcccaaa agagatgtgt ggctcagcat aagaggaaca ggagaaactg 19500
aggcctcttg ccctgagaat gagcttggaa gtggatgtcc cggcctcact caaaccttca 19560
gatgactgag gcccagccag gagcttgagt gtaccctcag gtcataccct gagccagaag 19620
cacccagcta atccactcct catcactgac tccctcccca taaaaaacct gtttgctgtt 19680
tcaggctgtt aagttgtggg ctgttttgtt acacagcaat ggataactaa cacacgaggc 19740
ctggcaagtg tggagcaaag ctgcccaagc cctcaagtct gttcatgtgg gtgttggcct 19800
gtgtttgcag aaatccagcc actgagtcct cccatgcagt cactactgcc ctctgcacag 19860
acacctgcca catccctgcc tgggccagga gctccactag tgcaggaatg gggtctgccg 19920
tcccaggagg atccctgaca cctagcacag ggctagcagc aggcagcact tggttagtga 19980
ataaactgcc cttcacctgt acacagaagg gatgtttcta taaggggtaa ttaagtacag 20040
agctgggaag ctatgctgac cagaaggctc taaaagcaat taaccaacga ggggaaaacc 20100
cttcctactc attctcggcc cattttattg agcactgacc atgtggaagg ccccctggtg 20160
agactgggga atgcaccaat aactgagaca gcttccggct gttgccctca ggatgcctga 20220
gctgggatag ggccagggtg gggtggtgc gtgtgacagg gttactgttc acaaccctgc 20280
cgggccataa gccctcccca acaattccaa aatccaaaac gctctgaaga tggaaagctt 20340
ttgttgctca tctggtgaca aaacctcatt tggtgcatgg gccgggtgcg gtggctcacg 20400
cctgtaatcc cagcactctg ggagccgagg ggaaggatcc cttgagctta ggagtttgag 20460
accagcctga gcaacatgtg agaccccgtc tctaccaaaa atacaaaaat tagccaggtg 20520
tggtggcgca ctcctgtagt cccagctact cgggaggctg aggcgggagg atcgcttgag 20580
cctgggaggt gggggctgca gtgagctgag attatgacat tgcactccag cctgggtgaa 20640
agagtgagac tctgtctcaa aaaaacaaag ttaaaaaaaa aaaaactgtg catgggtgtg 20700
ggctacagat agtcttttct gccctactta gaatgaacgt gccacatttg ctatagaaat 20760
attcaagggc tggtggcaaa tgccacacag accctgacgc tgttccaagt tctgagaagt 20820
cctgcattcc tcagggcccc agagtttcag agaagagtct gtaggcctga gttaagaagg 20880
aacgccttca aaagccctgg ggacaaaggg gaaaggggtg ccccaggact gcgtgggtac 20940
ctaccggaac gagccgtcca ggttggcacg gtggatgaag ctgagcttgg cgtcagccca 21000
gtagagcttc tgctcctcca ggtcgatggt cagtccattg ggccagtaaa tgtccgagtc 21060
cacaatgatc ttccgggtgc tgccatccat ccctgccccgc tcaatccggg gcgtctcacc 21120
ccagtctgtc cagtacatgt acctgtgacg ggggcagggc aagagaagca gctaacacag 21180
atctgttttt tgtttttgtc tgcatagatg cagacatgaa acaacagaca gtgaacttgc 21240
cctaaaatct cacccatcgg aaataaccaa caggtatggt ttcaggtatt cctgccttaa 21300
gctgggcaat caaaatatac tatttccaac ttgttctcag ttaacagtaa attctgggca 21360
ccttcccttc ttgtggatag aaagattcct tgttcttttg atgattgcct agtgtactct 21420
gctgtaagtt ttttaaagaa cttcaggtta tttctgattt ttttgctacc atgaaaatgc 21480
tgtaaatgaa cctctaaaag gcaattcaaa acactcagga tggaatatta tttagtggta 21540
taaagaaatg agctatcggc tgggcccagt ggctcacacc tctaatccca gcactttggg 21600
aggccaaggc gggtggatca cgaggtcggg agatcaagac catcctggct aacacagtga 21660
aaccccgtcc ctactaaaaa tacaaaacat tagccaggcg tggtagtgag cacctgtagt 21720
cccagctact taggaggctg aggcaggaga atcatttgaa cccgggaggg ggaggttgca 21780
gtgagcagaa atcgcaccat tgcactccat cctgggcgac agagcgagac tccatctcaa 21840
aaaaaaaaaa aagaaaagaa aagaaatgat ctatcaagcc atgaaaagac atggaggaaa 21900
cttaaatgca tgttagtagg tgaaagagcc aatctgtatg agtccagttc taaacactct 21960
ggaaaaagca aatacacaga gacagtaaag catcagtggt tgccaggagt tggagaggag 22020
agggatgaat gagtggagca cagaaaatca gggcagtgga actatcctgt atgacatgga 22080
atggtgggtg catgtcctta ctcatctgtc taaaccaaga atgtacaaat caagggcgaa 22140
ccctcgtgta aacgtggatt ttgggtgatg gtgcgtcagc cagctttcat cagttgtaac 22200
aaatgtacca ccctgcacag gatgctgaca gttgggaagg ctgtgtgggt gtgaggacag 22260

ggatgtatag gaactcagta cctgctgctc atcaattttg ctgtgaacct acaactgttt 22320
gaaaaaatta agtctattta aaaacaacaa aacatggcca ggcacgatgg cttgcacctg 22380
taattccagt acttcgggag gctgaggtgg gtgggtcact tgagccaccc tgggcaacat 22440
ggcaaaatcc cacctctaca aaaaataaaa attaaaaaaa agttagctgg gcatggtggc 22500
acactcttgt agtcccagct acttgggagg ctgacgtggg aggatccctt cagccctggg 22560
aggtcgaggc tgcagtgagc tgtgactgta ccactgcact ccagcctgga tgacagagtg 22620
agaccctgcc taaaaaaaaa aaaaaaaagg ctgggtgcgg tggctcatgc ctgtaattcc 22680
agcgctttgg gaggccgaga tgggcggatc acgaggtcag gagatcgaga ccatcctggc 22740
taacacggtg aaaccccgtc tctactaaaa gtacaaaaaa aaaaattagc cgggcatggt 22800
ggcggacacc tgtagtcaca gctactcggg aggctgaggc aggagaatgg cgtgaacccg 22860
ggaggcggag cttgcagtga gccaagatca caccactgca ctctcagcct gggagacagc 22920
aacactccgt ctcaaaaaaa aaagaataaa acccatggct gggatggacc ctgaacctgc 22980
agctgcagct gttcctgggt aggtctgtgg gcgacgtggc tttgcttctc catgttccca 23040
agagacaagc atcacccatc catgagaaac aagcacatcc tcagggcgcc cttacgtgat 23100
ctctggccaa tgaaccaaga caaagtgagc agacaccagg tctgggatgg caggtcccac 23160
ccccaccagt gcccagtgtg ccctgtttgg aggtgaccac agggtgtgtg cccagaggct 23220
gggcgtgact ctcagcggag accagagggg aaccacacca gcttggagga ctcagttccc 23280
atcccagcca gctgggatga gccacaggac acaagggctg gcagacctat tgtgttttgt 23340
ccacccttca cagcagagaa aggggacagt gcccagaatg tcctctgagg agcctcctcc 23400
cactcttggt ccttgtaaaa tggtgctgac tcccttgctc ccttcttcct ggggtgggcg 23460
gcaaacccca ttcccctcag ccttagcaag tgatttagaa acaggcagct cgcccaagcc 23520
aggcatgaga gtgatcccgg gacacaggga gaacaagccc cgctttgccc tctgggggtc 23580
tccattcagc agaagaggca aatgacagac acacagccgc ctcctccccc accatggtgc 23640
tctgcagcct caggagcctc aggtgcacca agggccaccc catccagggg gccatgcttc 23700
cttgagtggt atcgttcctg agcgagtacc atctccacct tccagagggg ctgtgacaag 23760
atcaacaaga atgagggcat aggagcctcg aaccaaacat gccctcttcc ctgcagaggc 23820
tgactgcgcc cagctgctat caccaagccc ctgctcctcc ggccccgtgg ggacagggta 23880
agaggggtgt cacatggaac agctctccaa acagtccctc tcaagctgct gtctcctgtg 23940
catctagtga gaacccaacc aacaaaggga aggtgggaat tgctattccc attaggcaga 24000
tgagaaaact gaggccccga aaggctggcc tgttccaggt tacaggcgct gagcggctgc 24060
tctgggaaca cacttggtgt ctgctgaggg cccgagcccg gccatcatat gactcaccct 24120
tcgccagcaa agcccgggtg tgggtgaact tttcctggca gcctgggact ccaaggtgct 24180
ggcagccagc ccagggaagg ctcccgcgtg cctgcggcag acgccttgct ttacctgcac 24240
gtccccaccc ctaggagcct ggacagagcc cagaccctcc gccacctcct gagaaggtat 24300
caggggcatc agtctggact tggggggggaa tccacacagg ccttccccaa atgctccacc 24360
gtggcccatg gaaaaggctg gaaaacgtgc aggagcagga gcctccgcat ggagcataat 24420
tcacattcct tccccgagtt tcataacaga ggcctgctgg tttccttaaa tggggaattt 24480
gcgagccagt cggtgaccag agactggttg gcgtggacgt gctcttgcag agtctcaaac 24540
gctaccacaa gcccagccaa attccacgga ggaaaatcga cttccgaaga aaagagctgc 24600
agcatggcct tcgtgcagag ccagctgcgg ttgtggttgt gtgttatttt agggaagggc 24660
cattttgcat tttaaagagg gggttgggtt tcaccctggc tttaatttga gacccggggg 24720
ccactgcagc cccttgtcag gctggtacag gccggggact cctcccatgc taagccagtg 24780
tctttctggc cccagatcct caggggccag agggtcatcc ccagagcccg ctctgccacc 24840
cacatgggta ccctgggcct gggagggatg tgccttccct caaccctgcc tggatgtccg 24900
cacggggcca cctgcattgc tgaaactgca acgaagtcga gtctcaggag gggccccccct 24960
ggctgcaggg ctcttgatcc ttttggccac gtgcacactg aggtggacgc tcggacccag 25020
agacccccctt catgatgatg gccggggcag gaaccccctc ctctgaggaa ggaccctggt 25080
gggggacagc actgcaggag ggcacaggag atgacggggg ctctagcagg gccgggagga 25140
aggccaagat gctcctcgca accgtgtgcc tgtggccagg acagaggaca aacccaccct 25200
ccactgtccc cactctcagg acagcagtcc tgccccagga ctcagcgccc acacttatgc 25260
ctgaggacca ctattcaagt cagtatttgg cgagcagggg ttgctgccgc gggcgctgtg 25320
acaggctgga atcctctccc tctccctctc cctctccgga gacatggagc ctacagggac 25380
agagtcagca cctcagggta ggaccatggc tggcgtcatc agcatcactg gatctgatga 25440
gtgggagccg gcatctcact gttttcactc tctcattcaa atgactggag caaagggaag 25500
gtgtggggag aggcccagga atcaacacta aggtcaactt tgcccccagg ggcaggggtg 25560
ggagtgaaca gccacaggtg tgatcctggg gagggcttct gggagagaat tcagaggcaa 25620
gcatgtagag gaaccatttc aaatagttaa gaaaagccag agccaaacag ggacagttgg 25680

ctcgcagaga tgatgcaggc aaagccagct cagatctgag catgggaaag actactccca 25740
accaagggcc cagcatctcc caaccaagca ccaagtacct cccaaccaaa tgccaagcac 25800
ctcccaatca aatacctccc aaccaagcac ctagcacctc tcaactggac accaactact 25860
cccaaccagg caccaagtac ctcccaacca agtgccaagc acctcccaac caagtaccaa 25920
ttacctccca accaagcgcc tagcacctcc caactgagca tcatgcacct cccaacagag 25980
catctagcac ctcccaactg atcacctccc aacctagcac cgagcacctc ccaaccaagt 26040
gcagagcacc tcccaaccaa gtgccaagca cctcccaatc aaatacctcc caaccaagca 26100
cctagcacct ctcaactgga caccaacaac tcccaaccaa gcgccaagca cctcctaaca 26160
aagtaccaat caccttccaa ccgagcacct agcacctccc aactgagcat catgcacctc 26220
ccaacaaatc acctagcacc tcccgactga tcacctccca acctagcact gagcacttcc 26280
caaccaacat agcaaaagcc ataaagaagt aaaaagacaa aaccacgtag gcatggagac 26340
tggacttctg gtggcgagga aagggcattt ttattataac gacagctaac atttgttgaa 26400
ctcacaaact gttcttggtg ttttcctcat gacatgcagc atggtcacgc ctctgtacag 26460
acaaggatac tgaggcacag agtggcaccg tgccaacctt gtctcatctt tttatcgaac 26520
ctacatgcag agtgccagca aatccagctg tcttttctct tcagaacaga tcccaaatct 26580
cgccactcct tacccccaca agtgaggtgt ccccgctgct gctttctgtc gccaggatcc 26640
cggtaataac cgtggagagg gctcctgccc ccacgccacc caccccacag ctcactctcg 26700
ctccagccac caggggatgc cttccagcac gagtcagagc tggcacctcc tctgctcgag 26760
acctcatgtg tcctctcctc acaccttggg ccctgtttcc ctacattctg ctacagcccc 26820
tcaaacaggc cccgccccaa accagcccag ggcctttgca ctggctgatc cctctgcctg 26880
gaccgcgctg cccccagaca gccacacggt tctcagcctc atctgcttcc agtctcgact 26940
caaaagtcac caagaggcct tcccagcacc tgagctccga cggaagcccc tcgccacagc 27000
acccaagcac tgctttatcc ccctacgcac acgtcccttt caaatactat tcatttacca 27060
tctcctccca ctcactgaaa gggccagaga ctgggctata cccgctgcgt ggggagcagg 27120
accaggcgca agggctcaca aatgcagtgg atgcctggtt gggaggtgag ggagctgcag 27180
cgacccacgc tgggagggaa cgcaatgaca ggaggagcgc aggtcctggc gacacgatgg 27240
ccatggcagc cgctggtgag caaccgcagg ccggccctgg gagagggctt ctagcaagct 27300
gctatcttca gcctctccga ctactgcaga tgccccctcc tagccagaga cactgctaca 27360
ccagccgacc cttccaaaaa gaaggtcagt aaccccgcga ctcctggagc cacagtgcag 27420
ggggagaggg ctgagagggc aacagttcac caagcggaac agaggctgcc ccggaggtca 27480
gctggctccc cggcagctgc aggggtggct agcccactcg gagggcagcg agggcatacg 27540
aggggctcca gggatgagtg gttgcccagc acagcacccc tgggaggccg ggggcacttc 27600
tcaggtagtg ggggcacgag gctgctctgg cctgacctca gggactcaaa atactttggc 27660
gataaattcc accgtgtccc acccctgctg gtaccccata cttacacaca gactggttca 27720
gatgcagaca ctctcgcgca catactcgct cacacgggca catacacgtg cacacacagt 27780
cacatgcgca cactcataca cacacaaata tccactcaca cgcatgcatg cacacacacg 27840
gacacacaca ggctcacacg tatgcacgca tatgcgtgca cacgcacaca cacacacaca 27900
cgctcacatc ctcccactcc cacactcagt tgctcagaca cacacacgcc tggctctcac 27960
acaaacctgt tgggctctga aaggctccag cccttcccat gctcgtcaga agccagtcaa 28020
tggcttccta agtcaccaca cagatcaaag aggtgaactt ggccacatgg cactctgctt 28080
cctgagctcc caaacaccag ccttggtgag gacagaccct caccccacac cctcattccc 28140
actaccctgg gcaggcccag aggaggggca tctgcaggat ctggcaacca gcccctcccg 28200
cccggctcct gcagccggca ccatgggagt caggggggagg tcactgcaaa gggcaacagc 28260
aagttggtgg ccccaggact agagcccagg ggtcttcagt cctactccag agcttggaca 28320
ctgtcccaca gggcatggcc aagggaaggg cttccagagc cctgacttca gggaggaggg 28380
caggcgggct cctgtggcag gcctggatgc atggccgccc actcctggga ctttctaacc 28440
tagaatatct aggtcaggct gggtgcagtg gctcacgcct gcaatcccaa cactttggga 28500
ggccgaggag ggtggatcac ttgaggttag gagtttgaga ccagcctggc caacatggcg 28560
aaaccctgtg tctactaaaa atacaaaacc tagccaggtg tggtagtgca cgcctgtaat 28620
cacagctact caggaggctg aggcaggaga atcacttgaa ctcgggaggt ggaggttgca 28680
gtgagctgag atcgtgccat tgcgcaaaga agatctaggc cggcccctca accggtgagg 28740
tccaggctgg gagtgctgag agactgtggt gacactgaat gaactaacag gcaaagggct 28800
tccaactgag cctgggggtg gtgggaaatg gctcttgtgt tctagtcaag acctctgcca 28860
accagttctg acactgaccc agcacagaac ctgacaggtc agcaagggcc agggcttagc 28920
acagcccagg taagggtgtg tgtacggccc ccagagtcac tcccaggctg caagaaaagg 28980
gacaaaggag ggacaagggg tggccaagca aactgttccc tctgctcggg agtctggggt 29040
gacctggcct agctggccag tggagctggg ccacctcccc ttaaactctc caccccggac 29100

ttcgactcca aagctttcct gccacccacg ctctccccac ctgggatcac ggccaggccc 29160
tgagccttca agggcccagg tgaactcagc cagactagga gctgaggagg acacagggca 29220
gcttccagaa cggacccgag aaccactccc agcaggttct gcttccagac aaggagctgc 29280
actttttcag ccaatgcaat tagaaagcca ggagaaggtg caaattccac ctgcctgagc 29340
gtccgcactt cccaggccgc ccaccataca cacagcaaag atgtgtttaa ccattcaaac 29400
ccatggccaa ccacatcggt tgcctcagac atgcaagttt taaaaaggaa cataactatg 29460
ggccaggcac ggtggttcac gtctgtaatc ccagcacttt gggaggccga ggtgggtgga 29520
tcacctgagg tcaggagttc gagaccagcc tagacaccat ggtgaaaccc catctgtacc 29580
aaaactacaa aaattagctg ggcgtggtgg tgggcgcctg taatcccagc tacttgggaa 29640
gctgaggcag gagaatcact tgaacccggg aggcgaaggt tgcagtgagc cgagattgtg 29700
ccactgcact ccagcctggg caacaaggga gactccatct caattaaaaa aaaaaaaaaa 29760
aaaaaggaac ataactatgg agtctcaagg ggaagtaatt ccttcaacaa taacaaatct 29820
tgaaagctga gctctttttt tttttgaga caggatctcc tcactttgtc gcccaggctg 29880
gagtgcagtg gtgggatcac agctcactgc agcctcgatc tcccaggctc aaatgatcct 29940
cctacctcag cctcccaaga agctgggatt acaggtgcat accatcacac ccgattcatt 30000
tttgtatact ttgaagagat ggggtctcac catgttgccc agtgtggtct tgaattcctg 30060
gactcaggtg atctgcccgc cttggcctcc cagagtgctg ggattacagg cctgagccaa 30120
caccccccacg ggttcatttt cagagtcgca ccgagtgctg gggttacagg cctgagccaa 30180
ccccccccacg ggttcatttt aagagtgaca ccgagtgctg gggttacagg cctgaaccaa 30240
ccccccccacg agttcatttt cagagtcgca ccgagtgctg gggttacagg cctgagccaa 30300
ccccccccacg ggttcatttt aagagtgaca ccgagtgctg gggttacagg cctgagccaa 30360
caccccccacg ggttcatttt cagagtcaca ccgagtgctg gggttacagg cctgagccaa 30420
ccccccccacg ggttcatttt cagagtcaca cccttttttct gaaaaacaac ttgggctcat 30480
gcaaattcga gagagagatg gtgacactcc ccgccccctg gacccaggtg gagtcgcagc 30540
agggtttacc cgtgagcggg gtccaaggcg atggccctcg gctggtcaag gtcctgccag 30600
aagagcacct tccgggatgt gccattgagg ttggccacct cgatgcggtt ggtctctgag 30660
tccgtccagt acagcttctt gcccacccag tcgcaggcga ggccgtcggg agagaccagg 30720
ccggagatga ccacgttctg cacggcggcc cccgtctggt tcaggtaggt ctgcttgatg 30780
gcctcctcgc tcacgtctgt ccagtacacg gctcccttgg aaaactggaa gtccactgcg 30840
gccgcatcct ccaggccgct gaccacgatg gtggactcca gcttgactcc gccggcgtcc 30900
accagccgta cgtcccggcg gttggcaaat agcaggagcg gcgaggctgt ggggcagaag 30960
caaaccgtga gggccactgg ctaagccagc aagatacaca gccctgggat ggagcactat 31020
gcccagagca ctcctggtac tgccctgccc atgcccaaga cctccagttc cttcctccca 31080
cccctaaggc gttgtcagga agttgcctgg gcagccccgg cccgcatcat tcagaggctc 31140
ctgcagcgca gcaaacagcc ttcttcccac attcggtgac agcacctgtt tgtttaccaa 31200
ctgttacgtc tgttccccca gatatgggtg acccttcctg ccatgcccaa aacctcccac 31260
atcgtcctcc agaggctaca ggggccctgt cctgttctgc agagaagcca catccccttt 31320
gttggcctga cacaggggat ggggacatgc aggcacagca ctggccatgc tgctcgctac 31380
agacccagcc acagggccac attttttgag gggttcagag cccaggccag acagagcctc 31440
aagattccct tacaagtctt tgaccactgt ccaagctcag gcccgtttcc ttggccgtgg 31500
catcagcttc ccatccaccc ctgtattcca tgtttctccc accctgcttc tggacattcc 31560
tacatttaaa gggtcactct ggaatgccac cccttggctc agacaccttc cacagctccc 31620
tgtgccagtg ccatgcagaa caaggtcaga ccccctagcc tggcctccaa ggccttggcc 31680
tctggcctca cctacacttc tctccaccac ccccaccccaa gcattcctga tctgcctgcg 31740
gccaggctgg ctccctcacc tccctgtgca ccgcagccct cagccccttc tgcctgtgca 31800
agaagcctca tctcacagac aacggtctca ttcccacaac gggctcaatg agaaatcagg 31860
agaggccttc agaccatcac cccaccagac acctcagacg tcggaccagg agggtccagc 31920
aaccccccaac acagactcag agggactaag aagccacatg aggagtgaac acaagatgtg 31980
gacaggagga ggttaagggc ctccagggag ctccatcagt ccgtgttctg ctgtcagcag 32040
ggttaggctg ggctggccac aaacaccccc aaaaaacatc tgaagccttg gcttgaaaca 32100
gctgacattc ctcatgaaaa ctgcagaccc ctgggtcctc ctgcgcagat gggggagccc 32160
agccaacccc acactcccac cttcaccaag aaagagaaag ccaaaacaaa ctcaactcag 32220
ccaatgacaa tcacagaact gaatcctgta gttagttcag ttggtttcat ttcagcaggg 32280
gaaagatttg cagcctctat gagggtagct gggaacacaa agggccagag catggcccag 32340
gagaccccag cgcagtgggg tagatggttc cgagcacagg cctccctgcc aagacaagca 32400
ctggctcaaa tcctggcccc tcccattccc aggagacatg ctccacagga tgggaggaca 32460
cacagaggac ctgaggccag gaaaatgaca gcggcgcctc cgccgcccca cccgtgctgt 32520

```
catcatctta ggtctacagt tctttgtggc aacgagggac actgtgaaag tcaaacaaca    32580
ggaaggcata ggccacaaat aaagacaaac gggacttcat gggaagctaa agattttgtg    32640
catcaaaaga cactatcgag agagtaaaaa ggcaacccac agaatgagag aaaatatttc    32700
caaatcatag atctactaag agattaatat ccatgaaata cagagaactc ctaaaactca    32760
acaatgagaa aacaactaag ccaactcaaa aatgggcaaa caacttgaac agacatttct    32820
ccaaagatga catataaatg gccaataaac acatcaaaac aggcttaata tatccctaat    32880
catcagggaa atgcaaatca aaactacaat aagataccat cttgcaccaa ttaggacggc    32940
tactatcaaa aaaacaaaat agcaagtgtt ggtgaggatc tggagcaact ggaacccttg    33000
tgcaccactg gcaaaaatgt gaaatggtgc agctactatg gaaaacagca tggcagttcc    33060
ccaaaaactt aaacacagaa ttaccatatg acccagcaat ttcgctttgg gttatatacc    33120
caaaagaact gaaaacaggg acacaatcag atatgcatac accttggatc acagcagcat    33180
ccttcccaac agctaaaaca tggaggcagc caggcatggt ggctcacgcc tgtaatccca    33240
gcactttggg aggctgaggc gggtggatca cctgaggtca ggagttcgag accagcctgg    33300
ccaacatggt gaaaccccgt ctctactaaa atacaaaaat tagctgggcg tagtgacggg    33360
cacctgtaat cccagctact cacaagtctg aggcaggaga atcacttgaa ccctggaagt    33420
ggacgttgca gtgagccaag attgcgccac tgcattccag cctgggtgac acagcgagac    33480
tctgtctcaa aaaacagcaa aacaaaaaca aaaaaacaaa caaacatgga agcaacccaa    33540
gcgtccctct actgagggat gaatagcggg gcaaaatctg ctccatccac acaatggagt    33600
actattcagt ctcaaaaagg aaaaagattc tggtcaggca cggtggctca tgcctgtaat    33660
cccagcactt ggggaggctg aggcgggtgg atcacctgaa gtcaggaatt caaggcccgc    33720
ctggccaaga ctggcaccna gctacacana aagtatangg ccccggaaa                33769


<210> 9
<211> 72049
<212> DNA
<213> Homo sapiens


<220>
<221> unsure
<222> (8356),(8385),(38585)
<223> Identity of nucleotide sequences at the above locations are unknown.


<400> 9
tataccttgc gcggaccttc ggctcctgtg gtgaagacaa tatgaagaaa atagaaatta    60
cccataattt tgccacacag acttagttgt gtccatgtat cttgtgcacc ttttttctgt    120
ttacggatca aaatcgactt ttagggtcag gcgcggtggc tcacacctgt aatcccaaca    180
ctttgggagg ctggagttgg ggttgggggg tggatcactg aagatcagga gtttgagacc    240
agcctggcca acatggcgaa actccatctc tactaaaaat aaaagattag ccaggcgtgg    300
tggtgggtgc ctctaatccc agctactccg gaggctgagg caggagaatc gcttgaaccc    360
aggagacaga ggttgcagtg agccaggatc acgccactgc actccagcct ggcaacagag    420
cgagactctg tctcaaaaaa aaaaataaaa ataaaataaa taaatacata aattgacttt    480
taggagattg gttcaaacaa tgtgtgtaat gttgtgtctg agtgtttttc atttatcgtt    540
catgcaaatt ccgacatcat tcactcttct ccagagtgtg ctgttttcct gcctgtgtca    600
tcacccgtca ccttgaatgc cctcgtttag gtaaaataag tacattttat tcaaaaatat    660
ttgaggacat ttgggttgtc tccaggttct tggtcttgag ttttgctgtt cttgtggagc    720
catggtggtg tctggttgca ggaacctcca tgcgttccag ctgctgcttc tgcctgtgtt    780
cttagagagg aaatgctggg gtccgcggtt cccgggctgc tgaccaggaa gcctgcggtg    840
ctttacggcc cttccagaag cgggagatgc ccccacttaa gtgtcagaca ggcctttcca    900
cctcactggc agctctgagc ggctcccttc tatttgcaga tgactgagaa gttaccaatt    960
tccacgttta ctgactgctg tttctcctgt taatttgtat ttatagtctt cgctaattta    1020
ttgctagggt tttggtgttg tccctattga cttgtatgcc ttttaatttt ttaaacaaca    1080
ttaatatact tcattttttt agagcagttt taagtttaca ggaaaattaa gggacaagta    1140
cagagagttc cttccacctg ctgtcctcct ctcctcctcc ccaccttccc tccttcccct    1200
attgtaactt tctttctgat attataaaag tcactcatgg ctgggcgtgg tggctcacgc    1260
```

ctgtaatccc agcacgttgg gaggcagagg caggcagatc acctgaggtc aggagttcca 1320
gaccagcctg gccaacatgg tgaaaccccg tctctactaa aaacacaaaa agttagccag 1380
gcgtggtggc gggcacctgt aatcccagct actcaggagg ctgaggcagg agaatggcgt 1440
gaacctggga ggcagaggtt acagtgagtc gagatcgcgc cactgcactc cagcctgggc 1500
aataagagtg aagcttcgtc tcaaaaacaa agtcacacac gcttcttgta cgagggtcat 1560
ttggccgagg ggccagatgg ctcaccatct agttgggaca ggccatgagc tcggaatgct 1620
ttttacatat ttacatggtt gagaagaaaa tcaggagaat aatgttttgg gacatgggaa 1680
aatgacatgg aatttgcatt ttagtgtcca taaatgaagt tttgtttgct cccagctgtg 1740
ttgactgagg caggctggct tcctacagct gcggcagagc tgaggaggcg ggaaggagac 1800
cgtgcaggcc gcagcaccga aaatatttgc tctctggccc ttcccagagt gcttgccgac 1860
ctctgtccga cagctagaag gaaggatagg acccgtccga cgataaccac tgttgacatt 1920
tgagcgcgtt tccttcccgg cttttgtgtg agagtggcag tctgtttgct tttgtggtcg 1980
ggatctgctg cacgcacggc gggctgtttg catgaggctt cctggaggat agggctgggc 2040
tcggagctgc acgcagtggg gcgtgtcctg catgcagtgg ggcctcagaa gagagctgtg 2100
gtgggcgggg cagtgccaac gctggtgggt gccaggcctc cacgctcaga tcagccccgg 2160
cgacaggttt gggccaccct ctctctggcc tctgtgcagt ggcccaggcc gtctgctctg 2220
cctggcacac ttgcctctgt ccttccactg aagcgctcct cttaccctct gctcccggct 2280
gggtacgttg aattgtgtcc ctcaaggaga tatgctaaag gtctaacccc aggaacctgt 2340
gtatgtgatc taatttggaa acagggtctt ggctgatgta atcaagcgag gatgaggtca 2400
ccctagagta gggggcctat atccacggtg ctggtgtcct catgagagca ggtgagcaga 2460
cactgacact caggggtgaa ggctgcatgg agtcagaaca gggcttagtg cgatggcggc 2520
cacaagccaa ggaactccaa gtatttcctg caacaccaga agctggaaga tgccaggaag 2580
gatcctgccc tggagccttc ggagggagtc tgtccctgca gacgtcttga cttttgattg 2640
cagggatgca tgtcttaggg tgtgtggggg ggtgcatttc tgatgttaga agccacctgg 2700
ttggtggcga tgtgtcacgg gagccctctg caggttctgc gtgtccatgt ggtcggggac 2760
agaggtgggc agggacggac ggtgtcgagc tggacatgtc catgacgtcg gccatccctt 2820
gggatggctt ttttgttttg aggataaggc tgcctgccag gaagctgtgc cctgcctggc 2880
ccttgcccca agcccctggc ctgtgcttgg cctcgcggaa gggatgtcgc ccttctctcc 2940
tgcatgcgtg cagggaggaa ggggagaggt cagcagcccg cctggaggag gctcgggcga 3000
ggggaaggtt tcactttcag gcaatgttgt ggggctgttt aaacaacccc aaagaaaacc 3060
atttggccaa actgttagtt tccaaacatt ttacttcctt ggtgtttaaa taaattccta 3120
ccaagactct gtagctggtc ccagggaagg agttggcctc tcttctttat agcccggcac 3180
agtcagtccc ctgcacctgc ccctcccaac cccaggcctg cttccccgtg gccatggctg 3240
ctgcccggac ctctctacac acagaacctc ctggaggcca gctgtgggca ccagccttgg 3300
cagggctgtg gcggagccca ggctgctggt actctctctg cagctgctcc ctgctggcct 3360
ggctggacag cgtccccacc accactgggg tcacctctgt gctggtcaca gctcactcag 3420
accttcaggc aaatgggttg gatcctgcct ctctcccagg tgtctcagtc tctgcaaaac 3480
tcaaaaacct cagaggcctt gcagcctgag gggtgtcaga gacacctcct tcgaatcagt 3540
aaacacctac agattcaccc cagcagtgaa aggactgctt cgccacagag gtttgattta 3600
ctcctaagta attggaaggg atgccgagaa taggttcctc atggtgggac tagaggccct 3660
ctgctgacct agttaacaga gggctagggc tgggtgtgct cagcccctga aggttctagg 3720
cccatttggg acaccccgcc agaacctgcc acaacctgcc atgtggtgac agctacctaa 3780
atcccagagg ctcttgagct ggagagcaga cctctcaatc tcagcaggcc ccccacacag 3840
accccataac cctagtctgc cttcacagta cagttcgtgg ctatgtgttc acggatggtg 3900
ttgttcacct aaggtctctg ccctgtgacc ccaagggcgt cctgagggca gattccaagt 3960
ctgtttcgtc caccccctcct tccctagcag cgggtccagg gcctggcctg aactagcttc 4020
ccacagagat actggtggga tgatgaaggc agccaggcgg caagtgaaaa acgcacttcc 4080
tgcatgtgct ggctcctggg attgaagtgt ttgaggaagc aaagtgaagt gagctttcct 4140
cttgcggctg tgtgtccttg ggccgggagc ctaccctctc tgagcgttgg ggtccttgtc 4200
agtagaatgg ggcatcctca tagctcaagg ggtggtgtgt gaaaattgtg ctattgtgtt 4260
actttaatga ttttttttt ttcgagacaa agtctcaccc caacgcgcag gctggagtgc 4320
agtggcgcga tctcagctca ttgcaacctc tgcctcctgg gttcaagtga ttctcctgcc 4380
tcagcctccc aagtagctgg aattacagga gtgcgccacc aggcccggca tatttttcta 4440
tttttagtag agaggggggtt ttaccatgtt ggctaggctg gtcttgaact cctgacctca 4500
ggtgatccac ctgcctcggc ctcccaaagt gctgggatta caagcatgag ccaccgcgcc 4560
cggcctactt tagtgatttc ttaggaggac agagggaacg ggctggcaag acaggcttgg 4620
aatgtgtttt gggatcaagt gccggtttct gtctggcact ggcgttctct gtggggccat 4680

```
gatggacaca ctgctgaggt caagcgtgat tcgtcttgcg ctgtgcctgg cagtctcatt     4740
ggaaagttct gtagacatcg tgtggatggg gctcttcccg gccaagccct tggggacctt     4800
ccaggactgt gatctcccca cagtggctgt taagcaggga cctttcgtga agtggagtct     4860
ctggtcccct ccaagtcata gctagacagg gactcgggca tcgccaagcc tggctgatta     4920
ttcactggat gaggagacag gcccagagag gggcaggaac ctgcccgagg tcacccagca     4980
ggccccagag gtttcggtct cggattctcc ctgctcatcc ctggatgtag tgctgctgtg     5040
gatgtggttc tgtgctgggg gctgtggaga gcaggggggct tgtgccagga ccccagtgag    5100
ggtggcgccc tcgccatgag gccgactgtt ggtatggggc ggccatccac tggggtgtgg     5160
ggaggaacag ctttcctgag gaggaggtgg cgggaggaac agcttccctg aggaggaggt     5220
ggcggtgctg tgtgacctgg gccttgaagg acaggtccat tgtcaacaga acattttggg     5280
agtggagcct agagggagaa aatttgttga aattcagatt cccctccccc taccaataca     5340
caccaaatca gatgcccctg accagatcta aatttggctc tcagagattt ccattgtagc     5400
tgggcacttg gggaaccttc taagtgctgc ctctgcctct ccccagcctg cctgcctcag     5460
tttccccagc cctgggcccg tgtcgctgtt gccatcacgt gggcgccctc tagtggagga     5520
atcagattat gcactccggg gcttggagca ggagtcagga ggggctcctg tctttccttg     5580
aaacgttgga tgccgggatc ctggaacagt ctctgcattc ctcctggcga gaaccagagc     5640
ctgggcacag gggaccatct gttgtttgaa ggctgcagcc tggcagggca ctcaggagat     5700
ctggcagttg gctgcagggc caggtctagg ggccagggca tcagggaggc tctgggctgg     5760
ttcagccccg ggcccctttg cagattgtga cctgggcccc tgtgcagggg catggccaca     5820
ggatgctggg aggggtctct gaccctgacc ttcttggctc tgtgcatcct tgagaccaga     5880
aaggtctgga acaaatgagt agacgatgcc ctaacctggg gagggagcca catcctgatc     5940
ccagcaacct cgggaaggat ctgtcaggat tatggggcac cctgggggcc ccaagtctgc     6000
atgggtctcc acttgcaatt tctgtaggaa gctctgataa atccaaactg gggtcctag      6060
gacacagtca gaaatgctga taccgttgtg tgtggagcct cgggccctgg gggtcaggag     6120
catgtggagg gtgggccacg gggggttcaga agagaatcct gtaacccccc accccccaaa    6180
ctgaagccca cttgagggcc atggctgaaa ggttgggggg tctccgtgcg tcctgtggag     6240
tgggtggtga ggagtccttg ggtttgcacg cctctgggcc tgagcggcgg gaccccgtcc     6300
acagcggatc cctgggccct gttgctcaga tgctctcaga gtgttgctgt ggccacggag     6360
ggagcctgag ttaagcttct cttgtgccgg ttgtacgctg tcaggtcaca ctggtgagtt     6420
aggcagggca cagatgccca gagcagaggg aactttcctt ggggattcaa cacgtgcaag     6480
tcttaggggc tggcaaatcc tgccctcagc tagagagggg gcttttattt gagaccagaa     6540
tcacctgagc atcctcctgt ccccagctgt gtccagcctg tctgcaggga catcctgaga     6600
ggaccaggct ctcccctcat ccacctgcct aagtgccact ctgaaccctg tccacctgtg     6660
ccgtggaggg gcgtgacctc aagctgctca gccagcagca ggcttggccc tgggggggcag    6720
cagagaccca ggtggctgtg gggtgggtgc ttcgtggcgt ggttctgaaa cttcgttgga     6780
agtgtgtgga cagtgccttg cctgttctct gtgggaccct atttagaaac gaggtctgag     6840
ttactggggg tcatcactgt gttctgatgg cccagctgtg tggaggccgc ggtgcagccc     6900
catccaagga gccagggccc tgggtctagc cgtgaccaga atgcatgccc cggaggtgtt     6960
tctcatctcg cacctgtgtt gcctggtgtg tcaagtggtc gtgaaactct gtgttagctc     7020
ttggtgttcc tgaaagtgcc cccgggtctc aggcctcaga accagggttt cccttcatct     7080
cggtggcctg ggagcatctg ggcagttgag caaagagggc gattcacttg aaggatgtgt     7140
ctggccctgc ctaggagccc cccggcacgg tgctggggcc tgaagctgcc ctcgggtggt     7200
ggagaggagg gagcgatgaa gtggcgtcga gctgggcagg aagggtgagc ccctgcaagg     7260
tgggcatgct ggggacgctg agcagcatgg ccagcagctg ggtctgcagc ctggtacccg     7320
gcgggacttg tggttggggc tggtttgtgg ccaggagagg ggctggcagg agacaagggg     7380
gactgtgagg cagctcccac ccagcagctg aagcccaatg gcctggctgt gtggctctca     7440
gctgcgtgca taacctctca gtgcttcagt tctctcattt gtaaaatgag gaaacaaaca     7500
gtgccagcct cccagaggtg tcatgaggat gaacgagtga ccatgtagca tgggctgggt     7560
gcgtgtcacc taacatcacc agcctttgca aggagagccc tgggggcctg gctgagtatt     7620
tcccttgccc ggcccacccc aggcctagac ttgtgcctgc tgcaggccct tgacccctga     7680
ccccattgca cctgtctcca caggagccga ggaggtgctg ctgctggccc ggcggacgga     7740
cctacggagg atctcgctgg acacgccgga cttcaccgac atcgtgctgc aggtggacga     7800
catccggcac gccattgcca tcgactacga cccgctagag ggctatgtct actggacaga     7860
tgacgaggtg cgggccatcc gcagggcgta cctggacggg tctggggcgc agacgctggt     7920
caacaccgag atcaacgacc ccgatggcat cgcggtcgac tgggtggccc gaaacctcta     7980
ctggaccgac acgggcacgg accgcatcga ggtgacgcgc ctcaacggca cctcccgcaa     8040
gatcctggtg tcggaggacc tggacgagcc ccgagccatc gcactgcacc ccgtgatggg     8100
```

```
gtaagacggg cgggggctgg ggcctggagc cagggccagg ccaagcacag gcgagaggga   8160
gattgacctg gacctgtcat tctgggacac tgtcttgcat cagaacccgg aggagggctt   8220
gttaaaacac cggcagctgg gccccacccc cagagcggtg attcaggagc tccagggcgg   8280
ggctgaagac ttgggtttct aacaagcacc ccagtggtcc ggtgctgctg ctgggtccat   8340
gcgtagaaag ccctgnaaac tggagggagc cctttgtccc cctgncttca gtttcctcat   8400
ctgtagaatg gaacggtcca tctgggtgat ttccaggatg acagtagtga cagtaagggc   8460
agcctctgtg acactgacca cagtacaggc caggcctctt tttttctttt tttttttgag   8520
atggagtctc actctgtcgc ccaggctgga gtgcagtggt gtgatctcag ctcactacaa   8580
cctctgcctc ctgggctcaa gtgattctcc tgcctcagcc tcctgagtag ctgggattac   8640
aggtgcctgc cactgtgctt ggctaatgtt tgtatttttg gtagagatgg ggtttcaccg   8700
tcttggccag gctggtcgca aactcctgac ctcaggtgat ccacctgcct cagcctccca   8760
aagtgctggg attacaggca tgagccacca cgcccggtca ggccaggcct cttttgaaca   8820
ctttgcacac catgggtctt ttcatccagg ggggtaggta cagttgtaca gttgaggaca   8880
ctgaagccca gagaggctca gggacttgcc cagggtcaca cagcaggatg tggcaggtgt   8940
ggggctgggc ctggcagcgt ggctccagct ttccagcata gaaatctgtg aaagcagata   9000
gtttgtcggt cggtagggga gactttctga gacccgcccc agcggctcag agggtagtag   9060
ccaggggcct tcctgggggc tcataaccca gaacactgaa tgggaaaacc ctgatggagg   9120
aggcgcagtg gagctgtggg tgccgatggg aagtcccaga ggagctggga ggtcagtagc   9180
ggtgctgccc tctgtggagc acttagtggg caccaggtgt gtttccaggt tcatggccct   9240
gggacctgaa gctcagaagg tgaagtaact tgcccagggc acccgtcggg cagcggcggg   9300
cagaggattt gtgggctgtg gagcctgtgc tcgtggccca gccctggggg ttgtgagtgt   9360
gctggccggg gagcttttcc tgcaagtgga ctggtgtcta ggagccagca tgtcaggcag   9420
caggcagcgg gagtgcagca ggcagcggga gcacagcagg cagagggcgg ggctcgagca   9480
gccatccgtg gaccctgggg cacggaggca tgtgggagag ggctgctcca tggcagtggc   9540
tgaagggctg ggttgtgccc cgaggagggt ggatgagggt aagaagtggg gtccccaggg   9600
gctttagcaa gaggaggccc aggaactggt tgccagctac agtgaaggga acacggccct   9660
gaggtcagga gcttggtcaa gtcactgtct acatgggcct cggtgtcctc atctgtgaaa   9720
aaggaaggga tggggaagct gactccaagg cccctcctag ccctggtttc atgagtctga   9780
ggatcccagg gacatgggct tggcagtctg acctgtgagg tcgtggggtc cagggagggg   9840
caccgagctg gaagcgggag gcagaggggc tggccggctg ggtcagacac agctgaagca   9900
gaggctgtga cttggggcct cagaaccttc accccctgagc tgccacccca ggatctgggt   9960
tccctccttg gggggcccca gggaacaagt cacctgtcct ttgcataggg gagcccttca   10020
gctatgtgca gaaggttctg ctctgcccct tcctccctct aggtgctcag ctcctccagc   10080
ccactagtca gatgtgaggc tgccccagac cctgggcagg gtcatttctg tccactgacc   10140
tttgggatgg gagatgagct cttggcccct gagagtccaa gggctggtgt ggtgaaaccc   10200
gcacagggtg gaagtgggca tccctgtccc aggggagccc ccagggactc tggtcactgg   10260
gcttgccgct ggcatgctca gtcctccagc acttactgac accagcatct actgacacca   10320
acatttacaa acaccgacat tgaccgacac cgacatttac cgacactgac atttaccaac   10380
actgtttacc aacactgaca tctactgaca ctggcatcta ccaacactga catttaccga   10440
cactgacatt taccaacact atttaccaac actgacatct actgacattg gcatctacca   10500
acaccaacat ttaccgacac caacatttac caacactgaa atttaccgac accgacattt   10560
accgacaccg tttaccaaca ccgacgttta ccgacaccga catttaccga cactgatatt   10620
taccaacact gacatctact gacgctggca tctactgaca ccgatgccag catctaccaa   10680
caccgacatt taccaacact gacatttacc aacactgaca tttaccgaca ttgacattta   10740
ctgacactga catctactga cactggcatc tactgacact gacgtttacc gacactagca   10800
tctactgaca ctgacattta ccaacaccag catctaccaa caccgacatt taccaacact   10860
gacatttact gacactgata tctactgaca ctggcatcta ctgacaccaa catttaccaa   10920
caccagcatc taccaacacc gacatttacc aacaccagca tttaccaaca ccgatgttta   10980
ccaacgccga cgtttaccga cgccagcatc taccaacact gacatttacc gacaccgaca   11040
tttaccgaca ctgacattta ctgacactga catctactga tactggcatc taccgacact   11100
gatatttacc aacgccagca tctactgaca ctgatgttta ccaacaccga catttacgag   11160
caccgacatt tactgacacc aatatttact gacatcaaca tttagccatg tgatgggggc   11220
cggcttgggg gcaggccttg ctcttggcac tggggatgct gcagagacca gacagactca   11280
tggggtcatg gacttctgct tcttctccag cctcatgtac tggacagact ggggagagaa   11340
ccctaaaatc gagtgtgcca acttggatgg gcaggagcgg cgtgtgctgg tcaatgcctc   11400

cctcgggtgg cccaacggcc tggccctgga cctgcaggag gggaagctct actggggaga   11460
```

```
cgccaagaca gacaagatcg aggtgaggct cctgtggaca tgtttgatcc aggaggccag     11520
gcccagccac cccctgcagc cagatgtacg tattggcgag gcaccgatgg gtgcctgtgc     11580
tctgctattt ggccacatgg aatgcttgag aaaatagtta caatactttc tgacaaaaac     11640
gccttgagag ggtagcgcta tacaacgtcc tgtggttacg taagatgtta tcattcggcc     11700
aggtgcctgt agacacagct acttggagac tgaggtggga ggatcgctgg agtccaagag     11760
tttgaggcca gcccgggcaa aggggacaca ggaatcctct gcactgcttt tgccacttac     11820
tgtgagattt aaattatttc acaatacaaa attaagacaa aaagttaatc acatatccac     11880
tgccctgctt aagacagaaa acatgggtgt tgttgaagcc agaggcagct gctggcctga     11940
gtttggtgat tggttcctaa gcagttgaag gcagttttgt tttttccatag atgtctgttc     12000
tccctttgct gggtgcagcc tcgccctgct gctgtggtcg ggtttcagtg gcctcgtccc     12060
gtggacgcag cctcgccctg ccgctgtggt cgggtttcag tggcctcgtc ccgtggacgc     12120
agcctcgccc tgccgctgtg gtcgggtttc agtggcctcg tcccgtggac gcagcctcgc     12180
cctgccgctg tggtcgggtt tcagtggcct cgtcccgtgg acgcagcctc gccctgccgc     12240
tgtggtcggg tttcagtggc ctcgtcctgt ggacgcagcc tcgccctgcc gctgtggtcg     12300
ggtttcagtg gcctcgtccc atgggcgtgc tttggcagct ttttgctcac ctgtggagcc     12360
tctcttgagc tttttttgttt gttgtttgtt tttgtttgat tttgtttgat tgtttgtttt     12420
tgttgtcgtt gttgttgccc aggctggagt gcagtggcgc gatctcagct cactgaaacc     12480
tctgcctcct tgggttcatg ccattctcct gcctcagcct cccacatagc tgggattaca     12540
agtgcccgcc accacgcctg gctaaatttt gtatttttag tagacagggg gtttcaccat     12600
gttggtcagg ctggtctgga actcctggtc tcacatgatc cacctgcctc ggcctcccaa     12660
agtgttggga ttacaggcgt gagccaccgc gcccagccct ctgttgagca tattttgagg     12720
ttctcttggt gccagtgata tgtacatgtg tccccatcgc accatcgtca cccattgagg     12780
tgacattggt gcctctcctc ggggtggatg cctccctctg tttccagcaa cttctgaagg     12840
attttcctga gctgcatcag tccttgttga cgtcaccatc ggggtcacct ttgctctcct     12900
cagggctccc aggggaggcc cgaatcaggc agcttgcagg gcagggcagg atggagaaca     12960
cgagtgtgtg tctgtgttgc aggatttcag accctgcttc tgagcgggag gagtttcagc     13020
accttcaggg tggggaaccc agggatgggg gaggctgagt ggacgccctt cccacgaaaa     13080
ccctaggagc tgcaggtgtg gccatttcct gctggagctc cttgtaaatg ttttgttttt     13140
ggcaaggccc atgtttgcgg gccgctgagg atgatttgcc ttcacgcatc cccgctaccc     13200
gtgggagcag gtcagggact cgcgtgtctg tggcacacca ggcctgtgac aggcgttgtt     13260
ccatgtactg tctcagcagt ggttttcttg agacagggtc tcgctcgctc acccaggcga     13320
gagtgcagtg gcgcaatcac ggctcgctgt agcctcaatc tccctgggct caggtgatcc     13380
tcctgcctca ccctctgagt agctgggact acagacacat accaccacac ccagctagtt     13440
tttgtgtatt ttttgtgggg ggagatgggg tttcgctgtg gtgcccaagc tgatctcaaa     13500
ctcctgaggc acaagcgatc cacctgcctc ggcctcccaa agtgctggga tgacaggcat     13560
cagccgtcac acgcagctca atgattttat tgtggtaaaa taaacatagc acaaaattga     13620
tgattttaac cattttaaag tgaacagttc aggctgggcg tggtggctta tgcttgtaat     13680
cccagtactt tgagaggctg aggtgggcag atcacctgag gtcaggagtt tgagaccagc     13740
ctggccaaca tgatgaaatc cagtctctac taaaaataca aaaattagcc gggcatggtg     13800
gcaggtgcct gtaatcccag ctactcggga ggctgaggca ggagaatcgc ttgagcccgg     13860
gaggtggagg ttgcagtgat ctgagatcat gccactgcac tccaatctgt gtgacagagc     13920
aagactctgt cttgaaaaat aaataaataa aaaaaatttt aaaaagtgaa caattcaggg     13980
catttagtat gaggacaatg tggtgcaggt atctctgcta ctatctactt ctagaacact     14040
ttcttctgcc ctgaaggaaa ccccatgccc accggcactc acgcccattc tcccctctct     14100
cccagcctct gtcaaccact aatctacttt ctgtctctgg gggttcactt cttctggacg     14160
ttttgtgtga ctggaatcct gcaatatgtg gtccctgcgt gtggcttctt tccatagcat     14220
tgtgttttcc agattcaccc acacattgtc gcacgttatc agaatctcat tcctgactgg     14280
gtgcagtggg ttaggcctgt aatcctaaca ttctgggagg ccaaggcggg acgatcactt     14340
gaggcaggag tttgagacca gcctggccag cctagcaaga ccccagctac caaaaaattt     14400
taaaagttaa ctgaacgtgg tggtggtggg cacttgtggt tcccagctac ctgggaggct     14460
gaggttggag gatcgcttaa gcccaggagg tcaaggctgc agtgagctat gatcgcacca     14520
ctgcactcca gcctggacaa cagagcaaga ccctgtctga aaaaaaaaac aaaaaaaaaa     14580
gttcctttct ttttgtggct ggatgacatc ccattgtatg gccacagcac attttgtttg     14640
tctgtttatc gggtggtggg cagtggtttc caccttttgt ctcctgtgaa taatgctgct     14700
gtgaacattt gaattcaagt ttttgtttga acacctgttg tgaattattt ggatatatgt     14760
gtaggggtag gattgctgag tcctatggta atgttaggtt tgacttactg aggaaccatt     14820
aaactgtttt caacagtggc tgcgccgttc tgcatcccca ccggcagtgt gtgagggttc     14880
```

tgactttacc tcctcacaaa cgcttctttt ccatttaaaa aaatattcag ccaggtgctc 14940
tggctcacgc ctgtaatccc agcactttgg gaggccgtgg cgggcggatc acctgaggtc 15000
aggagttcga gacgagcctg gccaacatgg tgtaacccca tctctaccaa aaatataaaa 15060
attagccggg tgtggcagcg ggcgcctgta atcccagcta cttgggaggc tgaggcagga 15120
gaatcacttg aacccgggag gcagaggttg cagtgagcca agatcgcgcc actacactcc 15180
agcctgggtg acaagagtga aactccatct aaaataaaac aaaaataaaa ataaataaaa 15240
atttattaaa acattcatca cagccagcct agtgggtgtc ccatgtggct ttgcctcgca 15300
tttccctgat aactaggatg ctgagcgtct tgtcccaggc ttgccacacc tcagcacttt 15360
gagatacgtc gcacagtccc catttgcgaa cgagaaatga ggtttaggga acagcagctg 15420
tgtcatgtca cacagcgagc aggggtctc tgagccgtct gaccccacag ccgaccaagc 15480
tccaatcctt accgcctcct agtgttgtgg atgtagccca gggtgctccc acatttttca 15540
gatgagaaca ccgaagctca aaacaggagc gttttgtcca cattggatac acgatgtctg 15600
tggtttggtc ctgaagtcac tttatatctc agtggtccag actggagtag gacagggggt 15660
tctggggaat ggggaaggtg tctcaggtga aaggaaggaa ttccagattc tccatactgt 15720
ccttgggaag ttagaagact cagagggtct ggcaaagtca gacaaagcaa gagaaatgca 15780
gtcaggagga agcggagctg tccaggaaca gggggtcgc aggagctcac ccccaggaac 15840
tacacttgct ggggccttcg tgtcacaatg acgtgagcac tgcgtgttga ttacccactt 15900
tttttttttt tttgaggtgg agtctcgctc tcttgcccag tctggagtgc agtggcacga 15960
tctcggctca ctgcaagctc tgcctcccgg gttcatgcca ttctcctgcc tcagcctccc 16020
gcgtagctgg gactacaggc gcctgccacc gcgcccggct aatttttgta tttttagtag 16080
agatgggatt tcactacatt agccaggatg gtctcgatct cctgacctca tgatccgccc 16140
gtctcggcct cccaaagtgc tgggattaca ggcgtgagcc accgcgcccg gcccgatttc 16200
ccactttaag aatctgtctg tacatcctca aagccctata cacagtgctg ggttgctata 16260
gggaatatga ggcttacagg ccatggtgct ggacacacag aagggacgga ggtcaggagg 16320
tagaagggcg gagagaggga acaggcggag gtcacatcct tggctttcaa aatgggccag 16380
ggagagacac cctctgagca tggtaggaca ggaaagcaag attggaacac attgagagca 16440
accgaggtgg ctgggcgtgg tggcttacgc ctgtaatccc aacactttgg aaagctgagg 16500
tgggtggatt gcttgaggcc aggagttcaa gaccagcctg gccaacatgg tgagaccccg 16560
tctctactaa aatacaaaa attagccagg cgtgatggtg catacctgta atcccagctg 16620
cttgggaggc tgaggcagga gaattgctta aacctgggag gcggaggttg cagtgagccg 16680
agatcccgcc actgcactcc agcctgggcc acagagtgag actccatctc aaaaaaaaaa 16740
aaaaaaaaga taaaaagacc aaccgaggaa ttgaagtggg ggggcgtcac agtagcagaa 16800
gggggatcgt ggagcaggcc accctgtggt catgcactgg aagctcatta cctgacgatt 16860
tggagctcat cactgggggc ctaaggagaa tagatactga aggatgagga gtgatggcgc 16920
ggggcacggg tgtctttggt ggccagaact tggggactgc tggggtgcct cactgcaggc 16980
cttctcagcg ccctttatat gcttacacag gctgtttcta agagggggat acattgcata 17040
agcgttttca gactacctca tcatgggtcc ctttctttac cctctgtggc cctggtggcg 17100
cactctctgg gaaggtgcag gtggatgccc agacccgccc tgccatccac ctgcacgtcc 17160
agagctgact tagcctcgag attgctgctg gcacctcctg ccccgggaca cctcggatgt 17220
gcccgtggag atgctggctc tgtgttttct gctggagttt ggtgcgtctt ttcctcctgc 17280
aagtggccac cgctcttggg tatgtcctca ggcttctgcg agtcatggct gcttctcagg 17340
tccttgccca gcgccaggag caaaccctcc tggcactttg ttcaggggtg gatgcgccag 17400
tgttcctgct gtggaccgcc atctcacatg agggtcttgg gcctgcaggc tcgttcagga 17460
aacacccgct gagtatgcag tgtgtgccag ctgtgtccca ggcaatggcg gggacagtgg 17520
ctgctgctgg ggttgtggtg gcttctgggg actctgggga cagctgaggt gcaaggagcc 17580
acggctcctt gaggatgcag ttggactcca ggtggaaggg atggttgggg gaggtataaa 17640
tggggtcagg gaggagacac atttggaaca atgggaacat ttttaagatg ctatgtcggg 17700
aggcaacaag gtggccaacc caggtgctga ggagcccaca ccagccctgg acgtgttttg 17760
ccgctcacct ttgctgggga gtggtgggag agaggattcc gttccacgtg gtggtgtgcg 17820
cagctgggct gtgtggagct gggcgctagg aggaaggtgc tttctgcggg gctagccggg 17880
ctctgccttt gaacacaatc aggctccagg ttttcagcat ccagtgcatg agaggacttc 17940
acgggcagct gtggctgatc ccttgatgaa ttgggagaag aacaaaggtc tatgaaatga 18000
ggtttcatgt agatggcatt agagacgccc acaacagatt tacagagtgg agcggagacg 18060
gcggatgggt ctgggaggcc cctcctgctg gccttgactg tgacagctgt cctgggaatc 18120
agcttccagg ccgccccagc agcctgactg acacacacag gggttttagc cccatcctgc 18180
gaccagctgt tgccatcatc agtgacagct gggagtggcg gtggttccag ccctgggcac 18240
cctccccacc tgctggggcc cacccagggc agtcctgaca cctacaggtt gcttggagcc 18300

gcatccgagt cctgccccac cacgtgtgaa gcccgagtgg tcgtgggctg aggtcccctg 18360
attgcatccc cacttccctt ctgcttcaca tagctgcctc ttctcaccgt ttttccagcc 18420
tcctgggcta ggaattccag tgttgtgctg gctttgcccc aggacacctc cttagccctc 18480
ttcctgagtc tagagccccg gggttggaa gtcctggccc ctgggacacc tgcagccaca 18540
ctcagcttct cctgtgagcc tccagcatgt cccctcagga ccaagccctc acgttcttgc 18600
ctccccgccc acctgggctc agccagggga aggcctggct gggagcgtct cccctctgcc 18660
ctgcccttct cccctcctac cctgcccttc tctcctctgc cccgccatgg cttttatatc 18720
ctgtgccaca agacatggct gtgtgtgaaa gtggcagggt ctggcatctc tgtgggtctc 18780
tgaggcccac gctccagtgc cactcttccc acccgctggc cgtgccctca tgctggaggg 18840
acagcccagc cctctcccga accccagccc catgtgccca gctgccccg gccctctccc 18900
ctggaagccg gggtcactcc agccgtatgc catggtgggg acatcctgct tccttggcct 18960
tccagggaag gtcctctttc caaatggcga cacctggtcc ctgcctggag gctggaagct 19020
gtggcccttg tatgcccctc cagggtctgt gcgctcggtt ggcccgagtt cccatcaccg 19080
tcatcatcac catcatcatt gtcatttcgc ttgtctgtga gccggcctgg tctcccagag 19140
cagagaccct ctgaggtcca gcctgagttg gggtctccgt gctgacccct gacggggact 19200
caggacgtac caggtctggg tcaggagtga cccccaaacc tcgtgccctt tgacaggcac 19260
ccctgacttt tgctaagtgg gtggaggtga catcacttac agcgggagtg atgggacagg 19320
gtctgttggc tgcactgtgc tcccagggat ctggggagag gctatatccc tgggctttgg 19380
cactgcagag ctgtgtgtgt ttgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt 19440
gtgtgtgttt gcgtgcgcgc acatgtgtat aagatctttt tttattacat gaagcaagat 19500
aactgttgct gtttcctttt gggttttgtg ttcaacagag tggggtactt cttccctcag 19560
acaacagaac tctccccttt aaacacgtgc tgtcagaggg tgggtcttgg gctcatgtct 19620
gtttgcacag ccgagtcaga ggaaacacag ggttcttcat aaaaacactg cacagcaggc 19680
gactgtccag agtcagcctg caggacggca gcagccctgc ccctcagagc acagctaggg 19740
tgggctgctt tgggatctcc cgtcattccc tcccagctgg cagccggcgg ccggcccatt 19800
ccttggtgtg ctggtcaggg gggcgtgcgc ctgctctgct caccctggga atgggacaga 19860
agctggcagc tcggagagga cagggctgga cccttgggtg gcctctggct ggaccatctc 19920
attgtcctca gacacagcct ctcgggtcta gtttcatttc ctgaaaaaca agtgcacaga 19980
actagagcag gagtcgagag ctacggcccc cgggccagat ccagccctgc cacctgtttt 20040
cacaccatgc tcaagctgag tgggttttac atttttaat tacttgaaaa aaaaaaagcc 20100
aaaggaggtt tcatgaccca tgaaaattat atggaattca aaaaaaaaaa attatatgga 20160
attcaaattt cagtgtccat aaataatttc ttgagacagg gtctcgctct gtcacccagg 20220
ctggagtgca gtgctatggc atggctcgct gtacccttga cctcccaggc tcaagcgatc 20280
ctcctgtctc agcctcctga gtagctggga ctacgggtgt gtgccaccaa gcccggctaa 20340
ttttttttta attttagtaa agacagggtc tttctatgtt gcccaggctt ttctggaact 20400
ccatcttggc ctcccaaagt gctgggatta caggctcgag ccacggagcc cagcctgttt 20460
ttgtttttc actgataaag ttttgccggg tgtggtagtg tgtgcctcta gcgatttggg 20520
aggctgaggt gggaggatcg cttaagccca ggagtttgag gctgggctca agtgatcagg 20580
aggtgaacta tgatcatgtc attgcattcc agcctgggtg acagagcaag aacctatctc 20640
ttaaaaatat atatttaaaa agtattgggt gtggtggctc acgcctgtgg tcccagctac 20700
ttaggcatct gaggtgggag gatggcttga gcccaggagt ttgaggttgc agcgagccaa 20760
gatcgtgtca ctacactcta gcctgggtga cagagcccag accctgcctc tttaaaaaaa 20820
aaaaccaaaa aacatgtatt ggaacacagc catgcctgtt cagtcacgtg ctctccatgc 20880
tgctttctgc tccagagacc cttatggcct gaaagctgaa aatattttct atcctttaca 20940
aaaaagtttg ctgacctctg tcctggaaaa ttcatctccc aagttctctt ccggcactgg 21000
cgttcctggg tgtcctaaat ttggcccctg ttatttctga actctgtttt ggctctgttc 21060
cctcccagga gccaggacag gcacgttctc tgcatcttgt cccctgacgc ccagaggctt 21120
ggctcggctc aggcattctt ggaaatatct ggctccagga aaggcagagg cctcctgagt 21180
cggcccagag ggaacctgcc ccaggtctgg gggaggcctg acccagcaga gtggcttttg 21240
ccgatgggtt gggccggtca agatgtgctg aaagttgtcc tcagaaggcc actttgggat 21300
tccttcctcc agtattagag caactgagag ctgctcattg caagcctgat gttttcccag 21360
ttggccgggt ccaccgggtg ccctgggatt ctgggatctg ggtggaaagt aggggcttg 21420
ggggagtgtc ctgggttctg gaatccaggt ggcaagtggt gaggttcagg gagtggcttc 21480
tgagccacca taggggtctc tgtgggaggc tctgcccatc caggagattc cgcaggccct 21540
gccggcccag agccagcgtc ttgcgcttgc cgaggctaca gccagcccca gccgggtgga 21600
acagcccgtc gcctcctctc actttgtttt ggggccacct gggagtgtgg agcaagggta 21660
gagagggagg aagtggctgc cggccgctgc ccagcaccct tgtttgcctt gggccctctg 21720

tgggctcctt tttattgctc ttcaatgaag ccagggaaat ggacttcctt gcctcacttc 21780
agttcaacat gtctggaagt ttggtattaa aattaagaaa gtgtggaaat agagcaagaa 21840
gagaaaaatc tctccaagag ataatagtga cctctgagct gggcgcggtg gctcacgcct 21900
gtaaatccca gtactttggg aggctgaggc gggcagatca cctgaggtcg ggagtttgtg 21960
accggcctga ccaagatgga gaaaccccgt ctctactaaa aataaataaa taaataaata 22020
aataaataca aaattagcca ggcatggtgg cgcctgccta taatcccagc taaggcagga 22080
gaatcgcttg aacctgggag gcaaaggttg cagtgagcca agatcacgcc attgcactct 22140
agtctgggca acaagagtga aactccgtct caaaaaaaat aaataaataa aaaataaaaa 22200
tagtgacctc tggccaggtg tggcagctca tacccgtaat cccagcactt tggaaggaag 22260
gccgagatgg gcagattgct ttagcacagg agtttgagac cagcctggcc aacatggtgg 22320
aaccccatct ctacaaaaat agaataaaat ttaagaggta atagtgacct tttggtagat 22380
cgaaacctgg attgctttct ttttctaaat gctgattctt ttctttgtgg tgtttgtgtt 22440
ctgtgccgat gtccctcccc cagccctgtt attgtgagtg gaagaagggg aaagggttcg 22500
cccgctactg tgagcccctc ctctcacgct gggtgtcctt ggagaagcct gcacttcttc 22560
attgtacgcc agggctgggt ccctccctgg agtggttctg tgctgctggg atggggccaa 22620
cccctcagat gttttctgag tgtcacacac aggtgtgtgc attcatggcc tttgcgtgtc 22680
ttcctgttgt ggaggcaaaa atgtgaagaa ccctagatga ttttgggacc agggctccat 22740
cacctgctgt tcattgcaca ccggagcatc caggcatggg tggagagctc agacttccag 22800
gcacggtcgc aggggctggt ctaaccatgt tcccgcccgc ctgctcgtca gaaccgcctg 22860
ttgggagctg ttatcatgat accatacctg ggccctgggc tatccgattc tgacttaatt 22920
gctccaggtt ggggccaggc cgttgtttgc tgttttgttg tttcttctgt gacgttagcc 22980
actgggctaa tctgagcccc tcagttacag gtggagaaac tgagacccat gggggtgcaa 23040
ggacttgccg aggacccaga gccccttggg ggcagagctg aggcggggcc tggctttggg 23100
tcccagagct tccagtcccc ttcccgctct cctaacagct tttttttttg agacaagatc 23160
tcaccctgtc acccaggctg gagtgcaatg gcatgatctc ggctcactgc aatcttcgct 23220
agctgcgttc cagcgattct cctgcctcag cctcccgagc agctgggatt acaggtgtgt 23280
gccgccatgc ccagctcgtt tttttttgta cttttagtag agatagggtt tcaccatgtt 23340
ggccaggctg atctcgaact cctgacctca aatgatccgc ctgcctcggc ctcccaaagt 23400
gctaggatta caggctggga tcacactgtg cctggcccta gcagctttgt cctgtgccat 23460
ccaacaacag atgaccgaag tctttgtttc ttaacatgca ttccatctgc cttacagttt 23520
tgccacctgc aaaacagagg acttgtcgct tttctggtaa gctggaaatg taatctggta 23580
gcaggaggcc tgtggaagct tgcctttaat ggccttgtgt ctctttcatc ctgtcctgag 23640
agccggagaa cttggatgtt gcacctaact caaccttcct gttaacatac agttctgcag 23700
gctcatggat catcagaacc acgtcctatc tcacgcggct gtatgcttcc gttggttcag 23760
gtgtttttac cttgacagta ttttctcctc ggtggctttt gcggtggttg cttttaatca 23820
gcattgactc ttcaagaaaa atatttagct gctacatctc agaggagaca gggtggaaag 23880
catctgagac ctgcaggctc agacttagaa ccagaagtgc cctcagagtt catccggccc 23940
tgacccagcg ggaaatgagt tcacagagaa gcgggagaac tttgccccag gccctgccgt 24000
tgctcataac tgccccaggt ccttacattt gctccaggtc ctgccccagg ccctgcagtt 24060
gctcataact gccccaggtc cttatatttg ctccaggtcc tgccccaggt cctgcagttg 24120
ctctgtgtgg tgggtgtgat ctggagccct ccgcccattg ctgcacctgg ggcaggcatt 24180
gctaattgat cccaggactc cttcctgcgg agcacgccct ggttctccag gcagccgctg 24240
cctgtcagcc tgcagtggtt cgggagagga cacctgcttg cctggtctgt tccaaatctt 24300
gcttctcatc ccagcacagg taggggggtgc tatgggaaag ggatcctcag ttggccctgt 24360
cactgctcta tcagctgggg acgtggcatc ctagtgaaaa catcatggcc gggcgcggtg 24420
gctcacgcct ggaatcccag cactttggga ggctgaggag ggtggatcac ttgaggtcag 24480
aagttcgaga ccagcctggt caacatggtg aaacccatct ctactaaaaa tacaaaaatt 24540
cgccaggtgt ggtggcgggt acctgtaatc cgagctactc gggaggctga ggcaggagaa 24600
tcgcttgaac ctgggaggtg gagcttgcag tgagccgaga tcttgccact gcactccagc 24660
ctgggcaaca gagtgagacg ctgtctcaaa atctcaaaca aacaaacaaa caaaaaacaa 24720
acaaacaaag cgtcatttat ccagcacccc tggggaacca tgctacctgg tgttttatgg 24780
tacctggcaa ggtgcaggtg aagttgctgc tcttgggcat tgaacccgtc ttgtttgggg 24840
cagctcaggc cccaggcagg gtccgggttg gctctcgttg gtgtggccct ggcccatcca 24900
gacctatatt tctgccgtcc tgcaggtgat caatgttgat gggacgaaga ggcggacccct 24960
cctggaggac aagctcccgc acattttcgg gttcacgctg ctgggggact tcatctactg 25020
gactgactgg cagcgccgca gcatcgagcg ggtgcacaag gtcaaggcca gccgggacgt 25080
catcattgac cagctgcccg acctgatggg gctcaaagct gtgaatgtgg ccaaggtcgt 25140

```
cggtgagtcc gggggtccc aagccatggc tcagccatgc agacttgcat gaggaggaag 25200
tgacgggtcc atgcctgggc ataagtgttg agctcaggtg ccccgacctg gggaagggca 25260
ggacaggaaa ggtgacagta tctggccaag gacagatggg aagggaccaa gggagctgat 25320
tagggagtgg ttatggacta ggaatgtcgg taacaatggt tagaaagtga ctaacatttg 25380
ttgagcacct gctgtgtgcc cggccctggc cgggagcctt cgtgcccaca gtgaccccgt 25440
ctgcaaatgt agttccttgc cctactcgca ctggggagca ggacgcagag ccgtgcaact 25500
cacaggtgcc aagctcagga ctccctcctg ggtctgcctg ggctgggctg tgcttgttgc 25560
ccctgtggcc cacgcatgtg caccttccac ctgaaagcca ggatcttcag gacgctcccc 25620
gaggaggtcg ttgtctggca caatgatttg tctcttcctg aaaaggtgac agagttacac 25680
tggagagagc agcatccagg tgcggcaggg acaggcctgg ggctcgcggg cagggactct 25740
gtgtcctgcc ggggtcccac actgcacctg cttgtcagag gcactcagtc aatctttgct 25800
gatgaaggat gagaggacag aggacgtgat gcttgctgct gcattgcctg cagtcctggg 25860
tgagatgccc gggttgactc tgctgcccgt cgggtggatg tgatgtcaga tccccggctt 25920
taaaatacga gggagctggg aattgaggga gcaggttggg gcagaaagca cagccccgtg 25980
gaagcctgga gctgaggcag tgtgggcgac ccctggagca gtgagtgctt ccttcatggc 26040
cttcatcgca ccctgcagtc ctcatgtagg ggatgccatc catgaattta gttttcccag 26100
cctcctttaa aaacgcgttc atgctggggc cggggcagtg cagtggctca catctgaaat 26160
cccaccactt tgggaggccg aggcgggtgg atcatgaggt caggagatcg agaccatcct 26220
ggctaacaag gtgaaacccc gtctctacta aaaatacaaa aaattagccg ggtgcggtgg 26280
cgggcgcctg tagtcccagc tactcgggag gctgaggcag gagaatggcg tgaacccggg 26340
aagcggagct tgcagtgagc cgagattgcg ccactgcagt ccgcagtccg gcctgggcga 26400
cagagcgaga ctccgtctca aaaaaaaaaa aaaaagtaca aaaaaaaaaa aattagtctg 26460
ggtgtggtat cacgcgccta taatctcact actcgagagg ctgaggcgga gaattgcttg 26520
aacccaggag gtagaggttg tagtgagccc gtatcgtacc actgccctcc acctgggcaa 26580
tagagcgaga ctctgtctca aaaagaaaaa aaaaaaaaga acatttatgc caggtgtggt 26640
ggctcatgcc tgaaatccca gaactttgga agactgaggc aggaggatca cttgagccca 26700
gaaatttgag agtgtcttcc ctgggcaaca tagagagacc tcatctctac cagaaaaaaa 26760
aaaattagcc cggcatggtg gcatatccct gtggtcccag ctacttaggg ggctgacgtg 26820
gcaggatcac ctgagtctgg aggcagaggt tgaagtgagc tgagatcatg ccactgcact 26880
ccagcctggg tgacagacag agaccctgtc tcaaaaaaaa aaaaaaaaaa aagcatttac 26940
tatccaccat ggaaggtgag actgacctgt gagtgattgt tcaaagaaca aaaaataaac 27000
cccagagata agacaaaagg gtgcctccat gggggtgtga tttaaagctg agaaattggg 27060
cttcttcccc ctcccctctc accccgtggt ttgctaaagg agatgggaaa aaggattctt 27120
tttttggctg aaatatttaa cactaaatta aagccaattt taacagcact ttggttgatg 27180
agtgaaatta acagactggc caaaaataaa cgaacggtct gtactatgtg aaaaagaggc 27240
agctttggcc atgctgggcc aatgtgagtt ttcagggttg ctgggaatgt ctgtgaatcg 27300
gaggaagggc ctagctggga ctctcaggag ccaaggccct gaggggcaac ttgcctggtc 27360
cctgccctga ggcgttcact gctttcttcc tgggccagat cacaggcccg gaggctggac 27420
cactgggctg gcactcttgc cgagctgctc cctgacttcc tgaccatgct cctttcagca 27480
gccttgctgc actttagttt ccttgaatga aaaatgggga tgagaatagc tcctacctcc 27540
aaggtgaatg gagtgagttc ggacaggtga ctccctggga ccagtgcctg gcgcctgaca 27600
aggtccagtc agagcccgca ctgctgttac tgataccctt ggctgtacca ggggagaact 27660
tggttgccat tgccaggtgt tctcccacca cccccactac tgtccctgtt tgatgtgtgg 27720
cgggaataaa gctgtgcaca ttggagcttt tggcacatcc tggctttcag gtgaaaggtg 27780
cgtgtgtgtt tgagggttta gcctggccaa cccagccatg aggtcggacc tgacctgggg 27840
gtgagtcctg agctcggcac ccctgagctg tgtggctcac ggcagcattc attgtgtggc 27900
ttgggccgca ccccttttccc tgctgggctg ttgatgttta gactggagcc tctgtgttcg 27960
cttccaggaa ccaacccgtg tgcggacagg aacgggggt gcagccacct gtgcttctgc 28020
acaccccacg caacccggtg tggctgcccc atcggcctgg agctgctgag tgacatgaag 28080
acctgcatcg tgcctgaggc cttcttggtc ttcaccagca gagccgccat ccacaggatc 28140
tccctcgaga ccaataacaa cgacgtggcc atcccgctca cgggcgtcaa ggaggcctca 28200
gccctggact ttgatgtgtc caacaaccac atctactgga cagacgtcag cctgaaggta 28260
gcgtgggcca gaacgtgcac acaggcagcc tttatgggaa aaccttgcct ctgttcctgc 28320
ctcaaaggct tcagacactt ttcttaaagc actatcgtat ttattgtaac gcagttcaag 28380
ctaatcaaat atgagcaagc ctatttaaaa aaaaaaaaga tgattataat gagcaagtcc 28440
ggtagacaca cataagggct tttgtgaaat gcttgtgtga atgtgaaata tttgttgtcc 28500
gttgagcttg acttcagaca ccccacccac tcccttgtcg gtgcccgttt gctcagcaga 28560
```

ctctttcttc atttatagtg caaatgtaaa catccaggac aaatacagga agactttttt 28620
tttttttttt tgagacagag tcttactctg ttgcccaggc tggagtaccg tagcgtgagc 28680
tcagctcact gcaacctccg cctcccaggt tcaagcgatt cttctgcctc agcctcctga 28740
gtagctggga ctacagacat gcaccaccac acccagctaa ttttttttat atttttagta 28800
gagacagggt ttcatcatgt tggccaggct ggtcttgaac tcctgacctc aggggaacag 28860
acggggttgg cctcccaaag ggcggaaata acaggggtga gccaccgttc ccggcctagg 28920
aaaacttttt gccttctaaa gaagagttta gcaaactagt ctgtgggctg gccttctgat 28980
tctgtaaaga aagtttgatt ggtggctggg tgcggtggct cacacctgta atcccagcac 29040
tttgggaggc cgaggtgggc agatcacctg aggtcgggag ttcgagacca gcctcaccaa 29100
cgtggagaaa ccccgtctct actaaaaata caaaaaaaaa attaaccggg catggcggcg 29160
cctgcctgta atcgcagcta ctcaggaggc tgaagcagga gaattgcttg aacctgggag 29220
gcggaggttg tggtgagctg agatggcacc attgcactcc agcctgggca acaaaagtga 29280
aactccgtct cagaaaaaaa aaagtttgat tggtgtaacc aaagcgcatt tgtttatgga 29340
ttgtctgtgg cagcttttgt tctgccgaga tgagttgtga cagatctgta tgggctctaa 29400
agcctaaaac atgtgccatc cgccccttta cagaaaaagt gtgctgacct ctgttctaaa 29460
gtattggaca actacaatgt ttgctcattt attattctat gatttgtttt ctgctttttg 29520
ttgttgttgt tgttgttgag atagggtttc cctctgtcac tcaggctgga gtgcagtggt 29580
gtaatttcag ctcactgcag cctcgacctc ctgggctcta gtgatcctct catctcagcc 29640
tccctagtag ctgggactac aggcacacac caccactcct ggctgatttt tttttttttt 29700
tttttttttt gtggagacag ggtttccgca tgttgcccag gctggtttca aactcctagg 29760
ctcaaacacc cacctcagcc tcccaaagtg ctgggattac aggcgtgagc caccatgccc 29820
agcctattct actgtttgta ttacatagct ttaaaagatt ttttatgact ttaagtcaca 29880
agggttcttt gtagaaaaaa atatatatat aggaaagtat aaaaagaaag taaaaattgt 29940
ccataacctc tccagccaga gacgaccgtt gctgacacct cagcatattg cctttaagtc 30000
tttttctct aagatagcat ttctcttcat cacagtcata tgctacgcag aattctgtat 30060
cctgattttt tcacttgaca ttacaacagg tatttgatgg cgctgtgaca aactctttgg 30120
cacaatcttt taaatgtatg aaatactcca ctgcacagat gtttgctttt aggcttaact 30180
gttcttttat tttgcgtgtg ctggttacag ccgggcacag tggctcatgc ctgtaatcac 30240
aacactttga gagggtgagg caggaggatc acttgagccc agaagtttga gaccggcctg 30300
ggcaacatag tgagacccca tctctacaaa aaactttttt aataagtcgg gcgtagtggt 30360
gcatagctgt agtcccagcc accaaggagg ctgagttggg aggattgctt gagccccagg 30420
aggttgatgc tgcagtgacc tgagattact ccactgtact ccaacctgag cgacagagca 30480
agacttgtct ggggaaaaaa aaaaaaaaaa tatatatata tatatatata tatatacata 30540
tatacataca cgcacacaca cataatataa aaatatatat ttataaatat ataatatata 30600
atataaaaat atatatttat aaataaaatt tataaattat atttataagt aaatatataa 30660
tatataatat aaaaatatat attatataat atataataaa atatataata taaaaatata 30720
tatttataaa taatatataa tacatactta taagtatata tttaaaatat atgtaatgta 30780
tattttttaa tgtatgatat ataatataca tttataaata cacatttata ttattttata 30840
taaaatatat ataaaatctc caagttgctt tttccaaaaa ggtgtcttgc tgcatttcaa 30900
acattcattt aaaaacttga atgctggtga tctggtccag aatgtgttca gtagctgctg 30960
ccagtggcca agcatctcgg gagatgtcta caaaacacgc tggttctggc ctggcgtggt 31020
ggctcacgcc tgtaatctca gcactttggg aggctgaggc aggtggatca actgaggtct 31080
ggatttcgag accagccttg ccagcttggt gaaaccccat ctctactaat aatacaaaaa 31140
aattagccag gcgtggtggc atgtgcctgt aatcccacct acttgggagg ctaaggctgg 31200
agaatcgctt gaacccaggg ggcagaggtt gcagtgagcc gagatcgcac cattgcactc 31260
caggctgggc aagaagagcg aaactccgtc tcaaaaaaaa aaaaaaagat gctggttcct 31320
aaaatgtggc ccttttcctc ctcacctgct gccagaccat cagccgcgcc ttcatgaacg 31380
ggagctcggt ggagcacgtg gtggagtttg gccttgacta ccccgagggc atggccgttg 31440
actggatggg caagaacctc tactgggccg acactgggac caacagaatc gaagtggcgc 31500
ggctggacgg gcagttccgg caagtcctcg tgtggaggga cttggacaac ccgaggtcgc 31560
tggccctgga tcccaccaag ggtaagtgt ttgcctgtcc cgtgcgtcct tgtgttcacc 31620
tcgtatgaga cagtgcgggg gtgccaactg ggcaaggtgg caggctgtcc gtgtggccct 31680
cagtgattag agctgtactg atgtcattag ccttgatggt ggccaggact ggtagggccc 31740
tcagaggtca tggagttcct tcgtggagcg ggtgctgagg ctgtatcagg cacagtgctg 31800
gctgctttca cctgggccgt ctcaccgaag tgtccatgga gcctgcgtag ggtgggtatc 31860
tgtgtcgatt ttacagatgc agaaacaggc tcagagaaac cgagtgactt ccctaaggtc 31920
acatacccag ttagagcaga gctgggccag gaagtgctgt ctcaggctcc tgaccaggtc 31980

tccttgcttt gcactcttgc caaaaccatg atccagaact gactttgagg tccccggacc 32040
tcaggctcct ccgaaatggc ctcttggagg ctgctgagcc acagcttagg acccacctcg 32100
agaggcaaat gtgctttgag ctgccaggcg tcctgggggc cctgccttgg gcacggggtt 32160
cagacaggcc ccagatgtgt ggggcgtctt tctggacttg agttttcttt tctgtgtggt 32220
ggacacagtg ctcacccctt aaagcacctg tgatgtgtgc agcagcccaa tccctgcctg 32280
tcgcctgttc tgctagggaa ggaaggaata cttcaggatg gcaggacaac agaaagaggt 32340
ccaggtttta gagcaagggc aggtcaaact tagaaaattc tggaatgagg atgtgcattt 32400
cctcttctgg atctgctaaa agaagaggga aggaggggct gctgggggag gagcccagag 32460
ccgagtttac atccggatcc cgcaaggcct cccctgccct gaggtcttgt tttgtgatgt 32520
gcttgtgtcc atcctggttt ctgccgtgtc cccaacatcc ggccaagctt aggtggatgt 32580
tccagcacac actcaccctg tctgtgcacc tgtttttgtg tccgtaagtg ggtatttact 32640
caccttacga gtgagccact gtgggaattc agggaggtgg cgcagtgacc acccctggag 32700
ggatatgtgt gtggcagggg tcgagggtct cgcccttccc tgcttcctgc gcgtggcttt 32760
ctccaggacg gggagggctg agctgaagag gtggggacag ttgcgtcccc ccgccaccca 32820
ctgtcctgcg gtgagagcag actcactgag cctgcccttc tcccttgtgc cttccagcta 32880
catctactgg accgagtggg gcggcaagcc gaggatcgtg cgggccttca tggacgggac 32940
caactgcatg acgctggtgg acaaggtggg ccgggccaac gacctcacca ttgactacgc 33000
tgaccagcgc ctctactgga ccgacctgga caccaacatg atcgagtcgt ccaacatgct 33060
gggtgagggc cgggctgggg ccttctggtc atggagggcg gggcagccgg gcgttggcca 33120
cctcccagcc tcgccgcacg taccctgtgg cctgcaagtt ccccaacctg gcaggagctg 33180
tggccacacc cacgactgcc cagcagcctc accctctgct gtgggagttg tccccgtcca 33240
cccctgggtg cctttgctgc agttatgtcg ggagaggctc tggtgacagc tgtttcctgt 33300
gcacctgctg ggcactaggt cccagctaat ccctgtgcca ggactctaat ttcaccctaa 33360
cacacatggt ggttttcatt gctggggaag ctgaggcctg agcacatgac ttgccttagg 33420
tcacatagct ggtgagttca ggatccccca gagataccag ggccagcact cgatccccac 33480
ccagccctga accccaccat gtgctgggat tgtgctggga gtgtccacac gcctgggacc 33540
ccagggctgg tgctctcatc tcctttttcc agatcatgag aatgaggctc agggaagttt 33600
gaaaaaaacc tatcccaagt cacacagcaa caggagcagg atttgaaccc agaaaagggg 33660
accgcacact ctgttctgct agagtagtta gctgtcctgg gtgatatggc aggtgacagg 33720
ggcaactgtg cttaacaaag gaaccccat cccccctgcc aagttgggag actagaaggt 33780
caggggcaga agctctgaag ggccaggtgc agtggctgac acctctaatc ccagcacttt 33840
gtgaggccaa ggcgggcaga tgatttgagc ccaggagttc aagatcagcc tgggtaatgt 33900
agtgagacgc catctctaca aaaaaatttt ttaaaaatta gctgggcatg gtggttcatg 33960
cctgtagtcc aagctacttg ggaggctcag gtgggaggat tgcttgagcc caggaggttg 34020
aggttgtggt gagctgtgat catgccactg cactccagcc tgggcaatag agtgagaccg 34080
tctccaaaaa aaaaaaaaga agaagaaaaa gaagctctga ggctccaagt ccccaggcac 34140
cccttggctt gagggcagac aagggaggag agggtcacct gggcagccct gacttttgtc 34200
ccctggcaaa gggaccttca gtgaccttgg ccctaggaga gcctctgagc acgtcagcca 34260
tgtcgaaccg ctcaggaagg gcagcaagaa tttggcttct gacctctgcc tctcctactc 34320
gccatctgca ctgggtgtgg ttgtgcccat tttacagatg aggaggctgg ggcatcgacc 34380
agctgaatgc cttgtcccag gtactgcgta ggcagagctg gcagttgaac cccgtgtcct 34440
ggttgtcgct gggggtgggc tgcaccctga cttgtgaggc cagtagcaag gtttgcacgt 34500
gacttcgtga ccgtcaccca gctctgcagc acatcccgtg acccagctca tccaggccgc 34560
atgcaaacct gttgccaggc gagaaaccag tcaccgcaca gctgtggttg cctgaaatga 34620
ttaagctcat taatcacccc ggagtgagga cagactcaga tgaaaaccag caaaagccct 34680
ggaaactcat gtgaccctgc caatgagggc ggccatgtgc attgcagcct ggccgtcact 34740
cctcggtacg tgttttggac ttaaacgctc cggatgttta ctgagtgctt gattaataac 34800
atggaaggcc tggtctcatt gctgtgggag tgaaggatgc acagccaggc ctgacatgat 34860
gagaacaaga acctggagtc tcgctgcctg ggtggtaatc ctggccctgc cacttagcaa 34920
ctgtgtgact gtagccaggt cacttaattt tgctagatcc tgcctgcgct tcagtggatc 34980
ttgctggttt tccaaggtgg ccaaacactt taaggcattc atgtggtcgc taggctgcag 35040
ggttgaaccc tggctcaccc cgcagggcgc cgtgtgctct gtggcctggc tgtgcctttg 35100
ctgacaccgt gcccgtgtgt gttcatgcag gtcaggagcg ggtcgtgatt gccgacgatc 35160
tcccgcaccc gttcggtctg acgcagtaca gcgattatat ctactggaca gactggaatc 35220
tgcacagcat tgagcgggcc gacaagacta gcggccggaa ccgcaccctc atccagggcc 35280
acctggactt cgtgatggac atcctggtgt tccactcctc ccgccaggat ggcctcaatg 35340
actgtatgca caacaacggg cagtgtgggc agctgtgcct tgccatcccc ggcggccacc 35400

gctgcggctg cgcctcacac tacaccctgg accccagcag ccgcaactgc agccgtaagt 35460
gcctcatggt cccccgcacc tcactccctc gttagatcag gctggttctg ggagctgacg 35520
ctgaaaggag cttctcatct ggggttcctg ggtgtacata gatggttggg taggttgtgc 35580
actgcacaag ctgcatgatg ctacctgggg gtccaggtcc aggctggatg gacttgttgc 35640
ttcatcagga catagataaa tggccaaaac tcctcagctg gaaggtcctg ggcaggatct 35700
ttgggtgtga aaaccagtca caggggaagg gtgcttgctc atactgccag cacagtgctg 35760
agtgctttcc atagcgctcg tttactcctc aagcctggag ggtggggagt agcatggtcc 35820
catttcacgt acaaggaacc cgatgcacag agaggtgtgg caacccatcc aaggccatac 35880
aactggggtg ggttgagccg gggttgactg tggcaggctg gctcaagagt ccctgctcct 35940
gaacccttgc caggcagcct ggcatcagct cggggaattt ttgccctgac ccttggaagc 36000
aagtgggcct ctttgttctc atgtcagtga tgagaagagt gactttccta tggcccctct 36060
ggagtacagg tgtttcctgt tggcgggctc ttcccccatg acatcagcag cgagctggtt 36120
atgattccct acgcagaact tgatagttta taaagctctt tgtcatccag gccccgttgg 36180
agtctcacgc agacctggtc gcaggcgggg ctggtcttgc ctgtcccagc tgcatggatg 36240
gggaacttga ggcttgcaaa ggttaagggg ctgttcgagg cccacgctgg caggagatgg 36300
gcctgggcca gagtctggga cttcccatgc ctgggctgtc tttggtcctg ttgctcacca 36360
tccctccctg gggccatgac cttagagagc caaatggagg tgcaggtaac ccacggcaag 36420
gaggggttgc catgactcag agtccccgtc ctgtggccgg cagtacctgg tgcaacgact 36480
tggatttcag accagccact gtagcccgct gacggtgcgc tcgaagtgcc acagcttctg 36540
aagccaggca ggactcaggc caggagactc tgttagctgt tgagagggag aggccaacgg 36600
atgttctggt tctgctagag agctggttct tcggatcctg gtaccagtgc actgagagga 36660
ggcccagctt gattctgggg ctgccttgtg gtggcatgtg ctgctcactg acaccctcga 36720
ggagtgtctt ctctcgggct tgttgactgt gcccggtttt ccgcagttca ctggtgcaca 36780
cataggcaca tagcaaaccg cacacacagt cgtgggtatg agtttcacta cattccacca 36840
ccagtgttca ctaccattac ctgccttccg tcttaagtgt tcatcattta aaaataaatt 36900
tattgggctg gacgcggtgg ctcatgactg ttatcccagc actttgggag gctgaggcgg 36960
gcagatcacc tgaggtcagg agttcaagac cagcctggcc aatatggtga aactccatct 37020
ctactaaaaa tacaaaatta gctgggcatg gtggggcatg cctataatcc cagctactca 37080
ggaggctgag gcaggagaat ggcgtgaacc cgagaggcag agcttacagt gagcccagat 37140
agcaccactg cagtccagcg tgggcaacag tgcgagactc catctcaaaa aaaaaataaa 37200
taaataaaag aaaaataaat ttatgatcta tttcaaaaat aacacatgta ctttgaaaca 37260
gcagagacac atatgacacg gagaatgaaa ttccccatag cgcaccccca agagacagcc 37320
ctggtccccc cgtctttccc gtggacctcc agcggggcag atgctgagcc gcctgttgtc 37380
gagtggcatg ctatcccgtc ctccagctcc tctgtggctt acagacaccc acctgcagcc 37440
ctgtctttgc ctcctctagc gcccaccacc ttcttgctgt tcagccagaa atctgccatc 37500
agtcggatga tcccggacga ccagcacagc ccggatctca tcctgcccct gcatggactg 37560
aggaacgtca aagccatcga ctatgaccca ctggacaagt tcatctactg ggtggatggg 37620
cgccagaaca tcaagcgagc caaggacgac gggacccagg caggtgccct gtgggaaggg 37680
tgcggggtgt gcttcccaag gcgctcctct tgctggtttc caggctgctg cccctgtcct 37740
tagcagaggg aggaaacaga ggatggctct gggtgaatga tgacttgggc ttcgattatg 37800
tagtcacagg gtatgaccct gagatgcgtg gaaccccgag actgtgatta tatgtagaaa 37860
ctgggtttcc ccgttgttta agtagtcatg gtggggtcag accccacagg acttttgtct 37920
tttcaagaaa gaaaatggtc gtgtgtcatg caggggtagt tggtactggt taatccaggt 37980
ttatcccttta ttttgtggga actgtacagt catttctgct acaatgctgt atatgctctt 38040
ctgaaagaca cctatgcaaa atcgcacagt aaaaatgaca caactcatag ggaaagcggg 38100
gccagggcac agccctcaaa atctccatca atgacatgta agaaaagaga ggaacctggg 38160
aaatagcaaa gtgccttttg cacattaaat ggttagctat atcccacaat actgtgcatt 38220
cgtaaacgtt aatgctgcaa taaatacggc acttcacctt gggaagatct ggagttggct 38280
tatgagtgtg gaagggtgta gcgcatgagt ttttgtgaaa cactggaagg aggattgtgg 38340
gaaatcaaat ggaaagttct caccccaggc gtggagaaga gtgggtcatg gccccagcag 38400
tgagcccagg gaggtcagag acggaggtgt gtgtgtgggt gtgaccctgc gcagttccct 38460
gccggctgta gtttttgca ttcgcttaat gtttctcgtg gaggaaattg tgcatgagca 38520
aatgtgaaac cgtgctgtgc tcaaattgtc ctaatacatc attgcattgg aacagattgg 38580
ctttntttt tttttttttt tttttttttt tttgaaatgg agtctcactc tgtcaccagc 38640
ctggagtgca gtggcatgat cttggctcac tgcaaccttt gcctcctatg ttcaagtgat 38700
tttcctgcct cagcctcctg agtaactggg attacagggc atgagccacc gcggccggcc 38760
agatttgcat ttttgaaaca actgctaggc tgggcgcggt ggctcacacc tgtaatccca 38820

gcactgtggg aggccgaggc aggtggatca cctgaggtca ggggttcgag accagcctgg 38880
ccaacatggt gaaaccccgt ctctactgaa tatacaaaaa tcagctgggt gtggtggcgg 38940
gtgcctgtaa tcccagctac tcaggaggct gaggcaggag aattgcttga acccaggagg 39000
cagaggttgc ggtgagccga gatcacacca ttgcactcca gcctgggcaa caagagcaaa 39060
actccatctc aaaaaataaa aaatagaaaa acaagtgctg tagcggaagt gagcactttg 39120
cggagtcagg cttgtgtggc ctgttccaca aatgatgtgc tcacggtggc ctcaggccca 39180
cctggagtct gcagcatggg gcacaacagg ttcattagtg tagaattcca ggacaggcct 39240
ggctcctaag cagccttctt ttacaaaaac tgcagagccc gcctgtatcg tagcactttg 39300
ggaggccgaa gtgggtggat cacgaggtca ggagttcaag accagcctgg ccaacatggt 39360
gaaaccccat ctctactaaa tatacgaaaa ttagctgggt gtggtggcac gcgcctgtag 39420
tcccagctac tcgggaggct gaggcagaat tgcttgaacc tgggaggtgg aggttgcagg 39480
gatctgagac catgtcattg cactccagcc tgggcaacag agcgagacgc catctcaaaa 39540
aaaaaaaacc tacagagcca cacggcctct ttctccaccg agtgttggtg tgggagcttg 39600
tgttattgtg gtgaaatctt ggtactttct tgaggcagag agaggctgag cgcctggaga 39660
gactttcaca tgggtcgcca tgtccgccgt cggtttcgct gttgtgctcc ccatctgaag 39720
gctggtgccg tccagacagg ctggacgccc ctttccacca gatccttcct cccgcagcag 39780
tttctagtta cgttgtactg tgaggtctgt gtccttggtt gatggcaaaa gtcagccgaa 39840
ttgaaattca gagccatgcc tggctccctg gagcttctct cctgggcagc tgtgatcatt 39900
gcctctgctg tggtgtgggt ggtggaaatg gattcctttc atcttgcttg ctacaggtga 39960
ctgtcacgtg gagtcctttg gagagaggga cgtgttaatt gatggatgtg gctcccatgc 40020
tgagaaagct cctgggcgta cattgcctta gagtttcatt ggagctgcgt tcttttatgg 40080
tgtctgctag gcagaagtga tgaagacttg gaagaaaacc cagaaggttt tccacttaat 40140
ttggaaaatg tgcttttccc ctcctgtgtc ttttgctaag gtccagcctc ctgcagcctc 40200
cccgctctgt ggactctggc tttgattctt tattaggagt ccccctgctc ccccaaaaga 40260
tggtgtctaa attatcatcc aattggccga ggttttgttt tctattaatt gtttttattt 40320
tttattgtgg taaatttata taacataaaa tttgccattt taattgtttt gttattgttg 40380
tttttgagac agggtctcac cccagtgccc aggctggagt gcagtggtgc gatcatggct 40440
cactgcagcc tcagcctcca gggctccagt gatcctctca cctcagcctc tctagtagcc 40500
gggactacag gcatacacta ccacatctgg ctgatttttt gtattttttt tttattgtag 40560
agacccgcta tgttgcccag gctggtctca actcctggac tcaagccatc ctcccacctc 40620
accctcccaa agtgctggga ttacaggcat gagccacaac acccagccat tttaattttt 40680
tttttttttt ttgagatgga gtctcactct atcgcccagg ctggagtgca gtggcgtggt 40740
atcaactcac tgcaacctct gcctcccagg ttcaagcgac tctcctgcct cagcctcctc 40800
ccgagtagct gggattacag gtgcccatca ctatgcctgg ctaatttttg tatttttttag 40860
cagagacggg gtttcaccat gttggccagg ctggtcttga actcctaacc tggtgatccg 40920
cccgcctcgg cctcccaaaa tgctgagatt acaggtgtga gccaccgtgc ccggcctttt 40980
tttgtttttg agacagggtc ttgccctgtc acccagactg gagtgcaatg gtgggctctt 41040
ggctcactgc agcctccgcc tcccaggctc aagttgtgca cctccacacc tggctaactg 41100
tattttatgt agagacagat ttcaccatgt tgcccaggct gggcttgaaa tggactcaag 41160
cagtccaccc acctcagcct cccaaagtgc tgagattaca ggcgcgagcc accgcaccca 41220
gcccatttta cctattctgc agttgacagt tcagtggcat tcagtcagtt cacgaggtaa 41280
ccatcactgc cattcatctc cagactactt caccttctcg gcagatgtcc gaaactgtcc 41340
gcattgaaca cactcctcat ctccctctga cagccaccat tctactttgt atctctctct 41400
gccttctcta ggtacctcat gtaagtggaa ttataccaat atttgccctt gtgtgactgg 41460
cttctttcat gtgacatggt gtcctcaagg ttcatctgtg ttatagcctg tgtcagaatt 41520
tccttcctta aagcctgaat aataacccgt tgtaaaggct gggcgcggtg gctcacaccc 41580
tctaatccca gcattttggg agtccgaggt gggcagatca cttgaggtca ggagtttgag 41640
accagcctgg ccaacatagt gaaaccctgg ctctactaaa agtacaaaat tagctgggtg 41700
tggtggcgcg cacctgtaat cccagttact caggaggctg aggcaggaga atcgcttgta 41760
cccgggaggc agaggttgca atgaaccaag attgtgcctc tgcagtccag cctgggtaac 41820
agagtgagac ttcctgtctc aaaaaaaaaa aaaatcatcg gatggatgga cggaccactt 41880
cttgttattt atccatccac gggtgctagg tttcttccac ctttggttgt cgtgaataag 41940
gccactatga acatttcctt ccgtggtgaa ggttttgtac tagtgaggaa aaggcgtgtt 42000
tgtggtgttg cataggattc tggtaagaaa gtttgcacta accataagta tttgtactac 42060
attaaaatga aagctcaggg gccgggcgcg gtggctcacg cctgtaatcc cagcactttg 42120

ggaggccagg gcgggcggat catgaggtca ggagatcaag accatcctgg ccaacatggt 42180

gaaaccccgt ctctactaaa aataccaaaa aactagccag gtgtggtggc gggcacctgt 42240
agtcccagct acttgggagg ctgaggcagg agaatggcgt gaacccggga ggcggagctt 42300
gcggtgagcc gagatcgctt cactgcactc gagcctgggc aacagagcaa gactccgtct 42360
cacgcaaaac tctgtctcac gcaagactcc gtctcaaaaa aaaaaagagt tcagggttta 42420
tgaaactggc cagccgcgta aagtttgctg tgttgttttt gtgcccggga ggagtgtggc 42480
cagggtgtca cgtcacacag tacacgtttc tcagatggtg gttctccaga ctgctgtccc 42540
aaagtctgtt tttgcatctg gttcccacag acccaccctc cacggtgagc ctgattttgg 42600
ccagggtagc tggaatcttg cttgtctttc agcccggcag ctgtaccagt ccagggtcca 42660
cagctagtgg cttttaggaa ggaatttgtt cagttggctt tgacacatgg ccccctaggg 42720
tccacagctc tgtagtgatg tggatgttgt tatctacaaa gacacatgat ccttcgtgtc 42780
cagatgaaag tgatgatgtc tttgcagctg cccagcaagg ctgtgtgtgt gtgtgtgtgt 42840
gtgtgtgtgt gtgtgtgtgg tgtgtgtgtg gtgtgtgtgt gtgtatgggg gagggaggca 42900
ccctttccat ctggggtgt gtgtgtgtgg ggtgtgtgtg tgtgtgtgcg cgtgtgtgtg 42960
gtgtgtggtg tgtgtgtgtg tatgggggag gcaccctttc catctgggtc caagagactg 43020
ggcctgggga agacgcttct ttttatctac ttagagactt tgttttattt gtattttttt 43080
gagacagggt ctcactctgt cacccaggct ggggtatggt gatatgagca tagctcactg 43140
cagcctcggc ctcccaggct gaagcgatcc tcccacctca gccttctgaa tagctgggac 43200
tgtaggcgtg cgtcaccata ctgagctatt gttttttttg tttggttggt ttaatttttt 43260
ttgatacaga tggagtcttg ctatgttgcc cagactagtc tcaaactcct gaactcaagt 43320
gattctccca cctcagtttc ccgacattct gggatcacag gtgtgagcca ctgctgtctc 43380
cctgttttat taactgctga aagacctaga taaagaaagt ctgaaaagac ttactatcag 43440
agcaccatcc taagatgatt ccctctgact caatggagag ggaggggagc ttttccttca 43500
ggcctgggtg gcaggagccc aggtgctcca ggccccattt gccccaggcc aaatcactcg 43560
ggaacttgga tgcagctgtc tttcagggta acccaaagga accagatccc cgcaggcagt 43620
aggcttctgg gctgtcctct cctcctacgt cagctcagta agagcccttc gaagggatgc 43680
tgtgtcggag gccccaaaag cccaggctca tccctgagat gcacagggtg ggctgggctt 43740
aggcagcgct cgagcatctc ctggacggtg accccagaga gtgtggagac ggagagtcct 43800
tgagagtcac tgagagacgt ggctgccctg ccttcccaag aggggctctg agtcattccc 43860
cacactcacc tgcccctacc caccctcacc tggcccccag cctcacctac ccccacatct 43920
gtaccgatcc ctttacccgc accttcccta cccaccctca cctcccctgt accttcacct 43980
cccccactca cccgcccctg caccctcacc tgtcccccac cttcacctaa cccccaccct 44040
cacctgccct cccctcacct ggcctccttc cgttggggaa ggggttgtaa ggggcggccc 44100
ccaaactgtc tgtcctggtg ccctgcagag aaaacagtac gtgagggccg cagtccaaaa 44160
gcttgagtcc tggaaggtgg aggagacagg gatgtgttgg gaagggcccc atggtcttgg 44220
atcccttctc gactgtcaat ggggccttca tgggagcgcc agtctagtga tgcacagctg 44280
ggtgcccggc gggtggctga ggaggcctaa agtccgaggc ggcaagagct cttccagagg 44340
ctgttgtcct aatcgctctg gcatactcag gcgggcacgt agttaggagc tgattggaga 44400
ggagagaccc ccacaccaat actgggattt gactttcagg ctaaacttga gaagtgtggc 44460
ctctgctgtc ctgccagagc tctccagcca gtgcccaggg ctctccagcc agtgcccggg 44520
ggtctccacc agtgcccggg ggtctccgcc agtgccaggg gtctccgcca gtgcccaggg 44580
gtctccgcca gtgctcagga gtcttggttt ctttgtctta cagccctttg ttttgacctc 44640
tctgagccaa ggccaaaacc cagacaggca gccccacgac ctcagcatcg acatctacag 44700
ccggacactg ttctggacgt gcgaggccac caataccatc aacgtccaca ggctgagcgg 44760
ggaagccatg gggggtggtgc tgcgtgggga ccgcgacaag cccagggcca tcgtcgtcaa 44820
cgcggagcga gggtaggagg ccaacgggtg ggtgggggtg ctgcccgtcc aggcgtgccc 44880
gccgtgtctt ctgccgaatg ccagcctctc acaggctggg gagactttcc accctgggga 44940
tccaatgggt ggctttccag ggtcccaaaa gcaaacacag gctctttcac agcccctcca 45000
ggaaagcaga aagccccaag ggctggaagg gaagggggag ctctgctgag aggttacaag 45060
gcagcgctgg ccgacgggag ttgcagttga taggttttgt atcatccttg ttaaacttga 45120
accctgtgca gaaatccctt ccacggcatg gggctgcct gttgactcgc tcctgttcca 45180
ccacagggag ctcctgggct tcttcctccc agaggccccc gacgctccca cctgttggtc 45240
gtcagagctt ctggttggtg ggaaggcacc caggaccttg aggtctccag agagaaaagc 45300
cagggaaaga gggagaccga aacccatgtg acatgaaact caggctccaa actgagcacg 45360
ggaacgtttg gggacaggag cgcgatggcc ttcctcagat agctgggggg ctggcatgaa 45420
gacgggagct acagccagca caggtcctgg gccgggagcc cagagattga gccctgactc 45480
tgtcacttac tggccacgtg accttgggcg ggtggcatag cctcttggag actcagtttc 45540
ctcattggta ggagtgacgg ccacagtggt gcggcctctg cagcacacgg ggggctcggt 45600

gggcggaagc cccgggtcta taaggcggct gtgcaggagc cagccgagct ggtctcccaa 45660
cagccagggc tccggggtcc ttagcagctg tggggggcct gcacctgttt cccatggctg 45720
ctgtcagaaa ttaccagaag ccaggtggct gagagtaatg gacacttgtt ctctcacagt 45780
tcctgagggc tgaagcccga gatcgaggtg tgggcagggc cctgcgccct ctgaaggctc 45840
tgagggaacc tttgggcttc tggtggctcc aggcacccct tgacttgtgg tcctgtcact 45900
ccagtctctc tgtctggctg cacatggcgt ggcctcttct gtaccattga aggacacttc 45960
agttggattt agggcctacc ctcacccatt gtggtcgtat cttgatcctt catgacattt 46020
gtaaagaccc tgcttccaaa taagctcaca ttctgaggtt ctggggtgag cgggaatttg 46080
gagagcattg ttcaactagt atagaatgtg acctgtcagc ctcgggcagc cctgagaggc 46140
aggggctttc cacagcccag ctgggtgccc tgggctccgt gctgtccgag gagacgccat 46200
ccccacaccc gtccttcacc cgccaccctc ccgcaggtac ctgtacttca ccaacatgca 46260
ggaccgggca gccaagatcg aacgcgcagc cctggacggc accgagcgcg aggtcctctt 46320
caccaccggc ctcatccgcc ctgtggccct ggtggtggac aacacactgg gcaagctgtt 46380
ctgggtggac gcggacctga agcgcattga gagctgtgac ctgtcaggta cgcgccccgg 46440
ggcctgccct aaccgcagac acccggcctt cattgtcagt aatggcagca gctgccacat 46500
tgtccgagac ctgccgtgag cccagtgccg cgccaggggc tttgtgtgta gcgtgttttg 46560
tcctcacact gacagctgta ggctggggtt ctgagtgagc cccacagggc agaggcagaa 46620
aatgagtctc agagagggtg agcgagctgc ttggggcccc acagcaggag atggagcagg 46680
actgcagcct agcctctgcc cccagcacct gcgcaagaag ctgctctgct ctggactgtg 46740
ttaggctgcg agggctggag agaaatgaga gttggtgctt agagaggggg cgcaggtccc 46800
catggctttt cctcttatga tgaggtagat gggtgaaggg aggggccatg cttgcagggg 46860
ccagtgaccg aggcccgccg ttggaactga tggccttcat cccgagccca gcccaggtgg 46920
gagcagggct ttccgagggc ttgtcttggg tcggcctgct tccagggact ctgctgcagc 46980
tcccacccct gtccaaagca tggaatcccc caggctccct ggcagtcctg tcaacctctg 47040
tcctcccaag ctgagtgtgg ggcaagttct ggaggtcagc actgctcagg ggggcccacg 47100
ggctgcttgc aggggccaac cgcctgaccc tggaggacgc caacatcgtg cagcctctgg 47160
gcctgaccat ccttggcaag catctctact ggatcgaccg ccagcagcag atgatcgagc 47220
gtgtggagaa gaccaccggg gacaagcgga ctcgcatcca gggccgtgtc gcccacctca 47280
ctggcatcca tgcagtggag gaagtcagcc tggaggagtt ctgtacgtgg gggctggcag 47340
tggggtgggc agggtggcct ctaaacccga cccctggagg aggctggagg ccagtgcaag 47400
atcctgtgtg gcctcagcca ggcggtggtc tctgccagat gccaactgtt gcccgctggg 47460
gttcagcgac atgtccgaat gtcccgaggc ctctgaggtt gttttctttt gccgcagaac 47520
aaatcaccac gaacagcgtt ttaagacaac accaactctt tttttttttt ttttttttga 47580
gtcaggatct tgctctgttg cccaggctgg ggtgccctgg tgcaaacaca gttcactgca 47640
gcctcgacct ctgggcttaa ttaagtgaac accttgcctc agcctcccag gtagctggga 47700
ctacaggtgg gcaccaccac acctggctaa tttttttttg tagagacggg gtttccccat 47760
gttgcccagg ctggtctgca actcctgggc acaagctatc tgcctgctgt ggcctcccaa 47820
agtgctagga ttataggtgt gagccactgg cctgacaaca cccacggatt gtctctcagt 47880
tctgtaaggc aaagtccagg cacagcgtgg ctcacctggg ttctctgctc agggtctcac 47940
ggggccagaa tcaaggtgtc aggaacgctg ggccctcagc ggaggctctg tggagaaatt 48000
agcttccttg ctcactcagc aggtagcagt tgtgggatcg aggttctgtt ttctctctgg 48060
ttattggtcg gggaccactc tcagctccta gaggccaccc caggtccttg ccccgtggcc 48120
ctctctgcct cagcagtggg ggctccctgc gtcagtccct cccgcacctt gagtctctct 48180
gatttgcttc taaagggccc tgtgattcgg ctcagccacc tttagattag gttagcctcc 48240
cctttgatag actccaagtc ggctgattaa taaccttact cacatctgca gaatcccttc 48300
tgccacataa ggtcatgacg ccgtgctggg gactggggtg ggaaattacg gggtcattta 48360
ggattctgcc tgccactgcc ttgctgtgtc ccagggcttg gggaggggc ctccacagct 48420
gggaccacag tccttcctcc cctccatggt aaccatctga ggattacttg agaccagcct 48480
gggcaacatg gtgagaaccc atccctacaa aaaatacaaa caaaaaggga ccaggctggg 48540
cttggtggct catgcctata atcccagcac tttgggagac caaggtgggc tgatcacttg 48600
aggttgggag ttcgagacca gcctgcccaa catagtgaaa tcccgtctct actaaaaata 48660
caaaaattag ctgggtgtgg tggcaggcgc ctgtattccc agctactggg gaggctgagg 48720
tgggagaatt acttgaacct gggaggcgga agttgcagtg agccaaaatt acgccactgc 48780
actccagcct aggcaataga gtgagactcc gtctcaaaaa aaaaaaaggg ccaggggtgg 48840
tagtgacaaa gagaccctat cccaaaaaaa ccgaacactg aatccttgag actgagtaag 48900
gacactgtga aatttttctg ggtggggcag ggaacagagc gtcttctgtc atttcttcca 48960
cctgggtgtg gtcagctctc cctccaagct gcctcctctt cttctcattg tccgggtgtt 49020

ggacacattt ggttaactgg atagaataac gcgagttccc agggacttgg tccatttgct 49080
attttatttt attttatttt attttatttt atttatttat ttatttattt atttatttat 49140
tgagatggag tttcgttttt gtcgcccagg ctggagtgca gtggcgcgat ctcggttcac 49200
tgcaacctct gcctcccagg ttcaagtgat tctcctacct cagccttcca agtaactggg 49260
attacaggca cccaccacca taccaggcta atttttttgt atttttagta gagacgggtt 49320
ttcgccattt tgcccaggct ggtcttcaac tcctagcctc aggtgatcca cgcacctcgg 49380
cctcccaaag tgctgggatt acaggcatga gccaccacgc ctggcaccat ttgctatttt 49440
aattcccatg tgtattagtg tcccacggct gctgtaacaa atgaccacaa actggatggc 49500
ttaaagcaac agaaatggat tcccccaatg tgctggagac cagaagcctg cgaccaaact 49560
gttgggaggg ctgtgcttcc tctgggggct ccagggagga tctatttgtt ggcccttcca 49620
gtgctgtggg tgccagcgtt ccacacttgt ggatgcgccg cctcaacctc tgcccatctt 49680
catgtgtcca tctcctttgt gtctgcgtct ttacctcttc ttcttgtctg tgttgcctct 49740
tataaggacg tttgtcattg ggtttagggc ccacccaaat catccgagat gacctcgtct 49800
tgagatcctt aacctgcaaa gacccttttt ccaaaaaaag gttatgctca cagattctag 49860
gccttaagac atgggtgtat ctttctgggg ggcactatcc aaccccttat acaatgaaag 49920
acgggaagag ggccaggtgt ggtagttcac gcctgtaatc tcagcacttt aggaagctga 49980
agcgggagga tcacttgagc ccaggagttt acaagtagct aggcaacatg atgagacccc 50040
atttctacaa aaagtaaaaa aaaaaaaaaa aaaaaaaaag ccaggtgtgg tggctcacac 50100
ctgtaatccc agcactttgg gaggctgagg caggcagatc acgaggtcag gagattgaga 50160
ccatcctggc taacacggtg aaaccccgtc tctactaaaa atacaaaaaa ttatggccgg 50220
gcgcagtggc tcccgcctgt aatcccagca ctttgggagg ccgaggtggg tgaattacaa 50280
ggtcaagaga tcgagaccat cttggctaac acggtgaaac cccatcaaga tcacaaggtc 50340
aagagatgga gaccatcctg gctaacacgg tgaaaccccg tctctactaa aaatacaaaa 50400
aattagccgg gcatggtagc gggcgcctgt agtcccagct gctcgggagg ctgaggcagg 50460
agaatggcgt gaacccggga ggcggagctt gcggtgagcc gagatcgctc catgccattg 50520
cactccagcc tgggtgacag agtgagactc cgtctcaaaa aaaaaaaaaa aaagaaaatt 50580
agccaggcac agtggcaggt gcctattgtc ccagctactt gggaggctaa ggcaggagaa 50640
tggcatgaac ccgggaggtg gagtttgcag tgagccgaga tcatgccact gcgctccagc 50700
ctgggcgata gagcaagact ctgtctcaaa aaaaaaagcc aggcatggtg gtgcatgcct 50760
gtagtcccag ctactcaaga ggctgaggca ggagggttgt tcgacccacg gagatcaagg 50820
ctacagtgag ccatgatcgc accactgccc tccagcctgg gtgacagagt gtgaccctgt 50880
ctcaaagtaa gtaaatagga ggagagacaa gtgggcagtt cagactgatg gtatgggcac 50940
agtagagact ggtgcagaca ggctggcctg tgatgtcaag caacttctgt aactgtttcc 51000
ggcatccatt tgtgtgtcaa tttccgtgtc agtaggaaga ctctgtaggc tgccaagagg 51060
aataagtggg aggatcctcc cagagaggcc gggcctgcag gagggccagt tctcatgagt 51120
tcttatttgg cccctaccct ccaggctgtg gttctgaggt gggagacaga gcctgacctc 51180
tgtttgtctt gttttgtctt tgcagcagcc cacccatgtg cccgtgacaa tggtggctgc 51240
tcccacatct gtattgccaa gggtgatggg acaccacggt gctcatgccc agtccacctc 51300
gtgctcctgc agaacctgct gacctgtgga ggtaggtgtg acctaggtgc tcctttgggg 51360
tgatggacag gtacctgatt ctctgcctgc taggctgctg cctggcatcc ttttaaaatc 51420
acagtccctg tggcatccag tttccaaagc tgattgtgtc ttcctttgcc ctcctttctt 51480
ttctactatg tgcattcggt gctatgaatt ttcctctaag tactgcgttt cctgcatctc 51540
acaaattttg ttacattttc attttcaggt agtttgaata tttttacact tctcctgaga 51600
tgacatcttt ggctcatgtg ttatttagaa gtgttgctta gtttctaaag agttggggct 51660
tttccagctg tctctctgca actgatttct aatttaattc tactgtagtc tgagagctta 51720
ttttatatga tttctgttat tttaaatgtg ttgggtgtgg tgtttttgtt gttattgttt 51780
ttgtgtcttt ttgttttgtt ttgcttcgtt tgttttgttt ttgagacagt gtcttgctct 51840
gtcactcagg ctggagtgca atggcgcgat ctcagctcac cgcaacctct gcctcccggg 51900
ttcaagtgat cctcttgcct cagcctcctg agtagctggg attacaggtg cacgccacca 51960
tacccagcta atttttgtat ttttagtaga gacggggttt caccatgttg gtcaggctgg 52020
tctcgaactc ctgacctcgt gatccgccca cctcggcctc ccaaagtgct gggattatag 52080
gcgtgagcca ctgtgcctgg ccattaggtg tgttttatca cccagcatca tgcagtttat 52140
cttggtgaat gttctgtgta ctcttgaaaa gaatgtggat tctgctgttg ttgggtggag 52200
tgttccagaa acatcaatta gatccagttg gttaatagtg ctcatcaggt tgtctctatc 52260
cttccttcct gactgcctgc ttgagctgtc agttattgac aggggtgtgg agtctccaac 52320
tctaatggtg gatttgttta tttctcctag tagttctatc tttttctctc cttctaccct 52380
tgatcctctt ctccccctag ggcttcctgg tgttggtggt gggagagtgg ggtagtgaag 52440

aacctggact ttagggccaa agaggccagg gttcaaatcc tggctctgtc acttcccagt 52500
tgagtgaccc tggctggtgc ctgaatctct gtgagcctcc acttcctcct ctgtgaaatt 52560
gagagcactt acctggcagg ctgtcatggg catcaagtaa cagggcactc cacctggacc 52620
ctgacacgtg atgcacagga atgccagctg ctatgccatg ggtgtggcag tagtaataaa 52680
gtgaccatct gtatcctcac cacagtgaag cctgtccagg gctttctctc ctatgccccc 52740
atgcctccag gtggccttgg atcctgttgg ttctgtgctc tgctcagcga cctttctccc 52800
gtgggagttc ctgggggttc agcttcatcc tacagacagc agcacacact ggctgtgcac 52860
cctttttttt tttttttttt ttttttttga gatggagtct cgctttttc gcgcaggctg 52920
aagtgcagtg gtgtgatctt ggctcactgc aacctctacc tcctgggttc aagtgatttt 52980
cctgcctcac cctcccaagt agctgggatt acaggctccc accaccacgc ccggctaatt 53040
tttgtatttt cagtagagat ggtgtttcac catgttggcc aggatggtct tgaactcctg 53100
acctcaggtg atccgcccac ctcagcctcc caaagtgcag ggattacagg cgtgagccac 53160
cacacccgga gtgccggttg tttttagcag tttgtcttgt tcctggagag actggctcct 53220
gcccaggagc tcggggagta gggccgcggg gtgctgcctc acacctcgag tttggccgta 53280
agcagagggg acattttgtg actgtccccc tcctgagctt cccagcagct tttctccaag 53340
ttacagccca aaagctcagg tggatttgca acccaacggt gtctgtgcac ctcccactga 53400
tgcccgaact gccctggcca agaaacgggg ccgtcagaac gctgcactaa ctgcagcctt 53460
gggcctccat gccagaggcc atgcccttcc atccaccacc ccctggcctg ggccctggcc 53520
ctcctggctc gggaactcca ggccccttcc tcacggatcg agagacgtgt atttaccgca 53580
caggtgcttg tcattctctt gtggcctctt ctccagggag atcacagaag gacagggcct 53640
cactgaggtc tcggacatgg accctttgat agtggcagga gccaggctgg gcaagaggcg 53700
gccacagtca cctcagcagt gccatcacca ccgccattca gcccttccct gagccgggcg 53760
cgcccctggc tctggcccca gtgtcccagt tacagctcac aggagcttgt ggtgcccagc 53820
ggctgcttct gattgagagt cgaggtcgga ggctttggga ggctgagagg ctgctcggtt 53880
tcacaactgc tgagggagac ttgggctcca tctcaggtct gccccatgtc gccctcaacc 53940
tccagccacc ggtcctccgt gtcccccatg gccaggcacg gcttgcagac atctgtcgtt 54000
ggctcctctc agccgtcgtg ggctgaccct ggcacgtcct cctgtggctg agcccagtgg 54060
ggacagctgc ttccttttat taccctagaa ctctcgtctt tgatcaggcc ccctccccta 54120
tgccacacag tccctgtcac tcgggtgagc ccagtagtca tggggaaggc ctgcgggttc 54180
caaacatcca aaggcttgcg tgcagcatga cagcttgaaa ccgatgtttt ttaccttgat 54240
cagatttcag cttggcgggg gctttgctca gctttcagtg aggcctgggc cgatttccca 54300
gcatcccctc ctgaggccag cctctgtttc ctgtgatttt ctgcacaaag tgggagggag 54360
gagtcttagg aaatgggggg ccacctcgaa acctaggcct cctctggctt ctctgtgcca 54420
gtgcccccac gctttgtgtc tgtgtcccca gcccatggga ctgtgttatt ccctgagtgc 54480
tgccgcatgc ccagcccgca ctgaggacgt ggagccccga ggggcaggat ggcctccatg 54540
gtcacacgta ggaagtggcc tccaccctcc gatgatcctc tcccccctc cctttcagcg 54600
ccttccccgg gggtgtcatc agccctcctg cctgtgcttt gtcccgtctt ctgcaggcgc 54660
atgggacgtg ctgacaggtc ctctgccggg ttcctgcctt gctatgcgca cgctggtcac 54720
cacagaggcc tggcccttct tctgtagcag tcccacaccc gcaacaggtg tggctgctga 54780
ccacctgctt tctgcccctc tggtcctgag gagggcgcag tgggcactca ggcgtggctg 54840
agcagatgtg tgttgccggg aggaggaagg actgctccag tcagggctga atttcccacc 54900
cggagcattt ctgctgtatt tggtgtagcg cctgctgctt aaagctctga ttcccagttg 54960
gcaccctttc ccttctgcat tgaaaaacat acggatgcat gtcttcttgc agtgaatgtg 55020
tattctccca gcctctcttc tgggttgggg ctggaggtgg agcggcacac aggagccgca 55080
gcgatggagg atgtgcgggt gcagcacccc gtacagcagg gatgccaaac ccgcgctgag 55140
tccctctcaa cttctgcttt gaagcccagt cacgccattg cctgggtttt gctgggcggg 55200
gctgcatgtg atgttctcct ctgtccctcc cccagagccg cccacctgct ccccggacca 55260
gtttgcatgt gccacagggg agatcgactg tatccccggg gcctggcgct gtgacggctt 55320
tcccgagtgc gatgaccaga gcgacgagga gggctgcccc gtgtgctccg ccgcccagtt 55380
cccctgcgcg cggggtcagt gtgtggacct gcgcctgcgc tgcgacggcg aggcagactg 55440
tcaggaccgc tcagacgagg tggactgtga cggtgaggcc ctccccgtca aggctctgcc 55500
aagaccctgg ccctgccctc cgggatacga gcttggggct gcctccggcc tcacaggagt 55560
aggggctctg aaaacctttg cttgcaggga gattgccaag tctgtctttt aggcccaaca 55620
aggaaaactc tgcagttcca cccatcctgt cccaccaggt agtgtggctt gaaggcagac 55680
tgtgagggtc tatctcacct tcctgcatta ggtcaggagt ttcacagaaa cctgaggcac 55740
attcaggggt gggctgcaga ggtccatggc tcacaccctg gaaaatccgc ccccaaaaga 55800
cagtgctgtc tccactgacc agtctgtggg atagtgctta agcctgagtg gtttctatca 55860

032796-132.ST25

```
acatgtagaa tcaggaggta taaagagatt tgctcaggca tcctgggccc tctctgacca     55920
gcaggatctt cctttagatc ttgacagtga aacacatctc ttctgtgccc cctgtgagtt     55980
ttctttcatt cattcattca ttcattcatt cattcattca ttcgagacag agtcttgctc     56040
tgtcacccag gctggagtgc cctggtgtaa tctcggctca ctgcaacctc tgcctccagg     56100
gttcaatcga ttctcctgcc tcagcctccc gagtagctgg gatgacaggt gcgcaccacc     56160
atgcctggct aatttttgta tttttagtag agacagggtt tcaccatgtt ggccaggctg     56220
gtctcgaact cctgacctca ggtgatccgc ccgcctcagc ctcccaaagt gctgggatta     56280
caggcatgag ccaccgcgcc cggcctgagt tttcctttta tgaaggacct gcttggttgg     56340
ttgcctgcca catgttgtca gcaccatggg cccaggactg ctgaggagct gttgatgccc     56400
tcgctctccc agagccaccg gctctgttag ataattcaca tgcagtctgg ccactgtcct     56460
acgtcctcat tcacaaagag cagacatttc gtagaagatg agggcctggg agtaacctcc     56520
ctgcatgttt ttctataaag gcatagtggt taagtccttc cagctcattg accattggag     56580
aattttatgg aggctgtaga ctaggggctg gtaaactaag ggcccagggg ccaaatccag     56640
cctgccacct acttttgtaa ataaagtttt cttggtgcac agccatgccc attcattcat     56700
ttgcacaatg tctgtggctg ctttcatgcc aaaagcagga gaactgagtg gttatgctgg     56760
agacctacgg ccttcaaagc cccagacctc acgtctggcc cttgacagac agagcttccc     56820
cagccctgct gcgcatcctg gcccagcatg tgctgtgtgt gtgatttcag cttgcaggag     56880
ccgtggttag gaattgtccc tgtgttggtc cattttgcat tgctatgaag gagcacctga     56940
ggccgggtag attatgaagg aaagaggtct gtctggctca tggttctgta ggcagcacca     57000
gtatggcacc cgcatctgct cagcttctag tgaggtctca ggaagctttg actcatggtg     57060
gaagtcgaag cgggagcagg tgcatcacat ggtgagagag ggagcaacgg agagagagag     57120
agagagagag agagcgcctc tccctcttgc cctcaccttg agaggagatg ccaggctcct     57180
ttaagtaacc agctcccatg tgaactcaca gtgagagccc atttgctact gcggagaggg     57240
caccaggcat ctgctcccat gacccaaaca ctgcccacca ggccctacct ccaaccttgg     57300
ggtcatattt tattctgttc tatgctatgc tatgctatgc catgccatgc catgccatgc     57360
tattcctatt ctattatttg agacagaatc tcgctctgtt gcccaggctg gagtgcagtg     57420
gcatgatctt ggctcactgc aacctccacc tcccaggttc aagcgattct cctgcctcag     57480
cctcccgagt agctgggatt acaggcacac accaccacac ccgggtaatt tttgtatttt     57540
caatagagat ggggtttcac catgttggcc aggctggtct caaactcctg gcctcaagtg     57600
atccacttac ctcggcctcc caaagtgcca tgattacaga tgtgagtcac tgcgcccagt     57660
gagggtcaca tttccgttga gatttggagg ggcagacgtt ggagccatct gagccccctc     57720
gtcccgctct agcttctcct cccgtgtgcc ccgcggtgct ggtggcaggc ccttacgccg     57780
gttctggctg cacgctctgt tccagaagct ttcttccctg cttggttacc agaaaatcat     57840
cccatccatt acaaggacag ggtcccctta tctcccattc ccagggcagg acaccggggg     57900
cagggcaggt ggggaactga gcaagttctc tgggggcagg cgtggctatg gctccctctg     57960
ggtgggcgtc tggggagggg tggaggcagc cgtcagcgcc ctggcttgct cttcctccct     58020
ggccagagac tgtggccttg tgctgctccc gtgtgggctg cctgcacctc cagtgggttg     58080
tgctccctcc cctcccctcc cctcaagctc tgctgagcac cactgccttc cacagccccc     58140
actctcggga ggcgaggctc ctcgtggcca ttcctgtcct tggcacccac ccccccacca     58200
acctggtaga gccttgggcg gggtctgtta ctccttgcat ggcgtagacc tccccacagt     58260
aggcacctga cacatacctc ctgggggcca ggcaggaggt gcgttgaggt ctcagccctg     58320
gcagtccctc ccctgcgtgg cataggcctc gccacagggt catcgagggt gggtggagac     58380
tgtactagac cactccccgc tggtcctaga aagggtccca tctgtctgct ctctgtttgg     58440
agtccagacc ttggttgctg tgccctgcat ggtgggctgg ggggcaccct ccagcctctc     58500
tgagtgcatg gcctctcctt gcagccatct gcctgcccaa ccagttccgg tgtgcgagcg     58560
gccagtgtgt cctcatcaaa cagcagtgcg actccttccc cgactgtatc gacggctccg     58620
acgagctcat gtgtggtgag ccagcttctg gcacggggaa ggggcgtccg ggctgggttc     58680
ccccaggaac gtggagttta ggggaggaga cgtgcctttc cagcggggct gggggctgtg     58740
tgggagactc aggcggctgg gaggctcctt gcgggaggca gggaagcctt tcccagggca     58800
gcggccagga ggacagactg tgagctgtgg gctcggcggc tacagagtct gcctcagtgg     58860
gcggggctga tggtgtccag gtgcctgcag cacgcaccca cccacgggac cttgctgagc     58920
agcgtctgtc aggcagcaag attacccgag ggctgcagtg gtcctgttcc ctggcagctt     58980
actgtctggc tgaggaggag tgatgttcac atatgcacac atgtcatgtg cacacacatg     59040
tacatgacaa catcccacat gctcctcaaa tagcatgacc tgtacagtca cggatatagg     59100
gcctagggga taggaggcca agacagtcag ggaagacttt ccagaggcag tggctcctga     59160
aaggctgtct gattcaggca ggaagggagc tgagttcaga taggaagtag caatgagtca     59220
ttgtgtctgg ggacatggcc actccttcgc tgcagaggga cctgggctga gagctcctct     59280
```

cttatggctg cagtcgggag agaagtctgt tggggggaga agggggcttc ctcaagggac 59340
tccctgtgcc ctttggcacc ttcgtgccag gtcaggcttg aggcctgaag gcagtggtgg 59400
gggccaccaa gggtcgcctc ctctgctggg caagttccca gtctgacggg cctgtgccgt 59460
gggccccagc tgtgggggcg ctgttgatgc gcagccaggc ctcgccgcca gagcccgcac 59520
gcttccattc cgctgacttc atcgacgccc tcaggatcgc tgggccggcc ctgtgggaga 59580
gtgaatgtgg cttttgccaa agttgagtct ggagcctgga aacttcccta tgggcagcct 59640
tgatagtgga gtggcccaag gagcccaccc agccgaccct gcccctcccg tggctggtgg 59700
gcggcaccag gggctgcctg gctttgctcg ttcaccaaca tcacccgggc tggccagggc 59760
gcgctcactt ctgccaccac cgagggccct gggcgaagga gtgaatacca ggctgccttg 59820
gcagggatgt gttgagggct gtggggagtc ggacagcggc gggggtcaga ggaggaggag 59880
ggtgcaccgt gcaggctgaa gggccacgtt accctgaggt tggccaggct ccccaggcct 59940
agcctcccag ctcccccact ttctccccac cctccaccag tggcaaagcc agccccttca 60000
gggcgcacgg tgtctgcccc caaggagggc ccattccgtt ggggttaatg ttggccacct 60060
ctttctgttt gtctctggca gaaatcacca agccgccctc agacgacagc ccggcccaca 60120
gcagtgccat cgggcccgtc attggcatca tcctctctct cttcgtcatg ggtggtgtct 60180
attttgtgtg ccagcgcgtg gtgtgccagc gctatgcggg ggccaacggg cccttcccgc 60240
acgagtatgt cagcgggacc ccgcacgtgc ccctcaattt catagccccg ggcggttccc 60300
agcatggccc cttcacaggt aaggagcctg agatatggaa tgatctggag gaggcaggag 60360
agtagtctgg gcagctttgg ggagtggagc agggatgtgc taccccaggc cctcttgcac 60420
atgtggcaga cattgctaat cgatcacagc attcagcctt tcccactgag cctgtgcttg 60480
gcatcagaat ccttcaacac agaggcctgc atggctgtag caacccaccc tttggcactg 60540
taggtgtgga gaaagctcct tggacttgac cttcatattc tagtaggaca tgtgctgtgt 60600
tgtccacaaa tcctcatgta ccctagaaat gaatgtgggg gcggctgggc tctctccaga 60660
gctgaaggaa tcactctgta ccatacagca gctttgtctt gagtgcagct gggatttgtg 60720
gctgagcagt tacaattcct acgtggccca ggcaccagga acgcaggctg tgtttgtaga 60780
tggctgggca gccgcaccgc agagctgcac catgctggtt tgtatcacat gggtgaccat 60840
ggtatgtcta agaaggtgga gtccctgtga ggtctgcagg tgcccccaca gctccaggcc 60900
accttgagga ttgcctctgc ctgcccagcc ctgagttccc tctcccctgt cctgtcccac 60960
tgtcacccca agccggcctc attgggagcc tgttggatgg cagggtatag atgtaacctg 61020
attctctctg gggagcgggg ttatctggct tctcaagagc tcctaggagc ccacagtggt 61080
ggcaccatca cagtcgcagc agcccccaga gaacgcggcc ctgtctgttc ctggcgtgct 61140
ctgtgctgcc ccgcctgggt tccctgcccc agtcgcaggc cccttggagg aggtaccatg 61200
tgtctcccgt ttcacagatg agccccgggg agctcactct agtagtggcc agagaggcct 61260
gcggctcagg gagcggggca catttccaac aggacacacc gccctggtct gagtctcgtg 61320
ggtagtggga gcagaggaga gcgccctatg tctgtggggc ggcttggctg agcctggaag 61380
ccacctgacc tcccccgtcc cttccctgcc aggcatcgca tgcggaaagt ccatgatgag 61440
ctccgtgagc ctgatggggg gccggggcgg ggtgcccctc tacgaccgga accacgtcac 61500
aggggcctcg tccagcagct cgtccagcac gaaggccacg ctgtacccgc cggtgagggg 61560
cggggccggg gaggggcggg gcgggatggg gctgtgggcc cctcccaccg tcagtgctgg 61620
ccaccggagg cttcccgggt tcctgggggc tgtgccaccg cctctgaggc atgcttgctt 61680
tcttcccttt tcaaacccct ctgcttcctt ctttaatgac attgttgatt gtggataatc 61740
tgaaaactac acaaaaatat aaagagccaa aatctcaccc aaatccacct cctagagtgg 61800
ctgttgggct ccgtcagcat ccaggcggcc gtctgtgttc cgcacggccc agcccatcga 61860
tagccgcctg caccaggcct gtctgccctc tgtgagcctc cccacagggt tccctccaca 61920
aacaccctgt tctcccaccc agggctggct gcttcctgga aaacagctgg atggttttgt 61980
gcatgacaga caaacacagg gtgattttcg tggctaaaat actccctgga gcttttggca 62040
gggtgagggg ctggctccag ctgagccacg ccttgagtga aatgactgtg aggagaataa 62100
actgccgctg ccctccagga tcactggggc tggctgggga gaacccccgt ttctgggagc 62160
acagtcccag gatgccaagg cgagcttggt gccgagatgt gaactcctga gtgtaaacag 62220
cgggggctga cttgacatgc tttgtatgct tttcatttgt tcctgcagct gtatgcccct 62280
aaggtgagtc cagcccccct ctgcttcctc tggggcctcg ccagtgagcc ccaccttgct 62340
ggggctggtt cctcctgccc ttctgggtat ccctcacatc tggggtcttg tcttcttgtt 62400
ttatttttct tttttttttg agacggagtt tcacttttgt tgcccaggct tcagtgcaat 62460
ggtgtgatct ctaggctcac cgcaacctct gcctcccagg ttcaagcagt tctcctgcct 62520
cagcctccct agtagctggg attacaggca tgtgccacca cgcccagcta attttgtatt 62580
tttagtagag atggggtttc tccatgttgg tcaggctgat cttgaactcc ctacctcagg 62640
tgatccgccc accttggcct cccaaagtgc tgggattaca ggcgtgagcc accgcacctg 62700

gcctttttct tttcttttct tttctttttt ctgagacagg gtctcgctct gtcacccagg 62760
ctggagtgca atggtgtcat catggctaac tgcagcctct accttctagg ctcaagcaat 62820
cctcccatct cagcccctaa gtagctagga ctgcacgcat gcatccccat gcccagctaa 62880
tatttacatt ttttgtagag atgaagtttc actatattgc ccaggctggt ctccaactcc 62940
tggactcgag cgatcctcct gcctcggcct ccccaggtgc tgggattaca ggcgtgagcc 63000
accgtgcctg gcctggggta ttgtcttctt atggcacctg actgtggtgg gccctgggaa 63060
ggaagtagca gaagagggtt cttcttggtt tcctggacag taactgagtg ttctggaggc 63120
cccagggcct ggctttgttt agggacaaag ggaactggta accagaagcc gagagtttaa 63180
acacccactg cccttcttcc ctgctcctgc tgctgcaacc cagcttaacc agccaggagt 63240
gctaggaacc caagcagggc ccccgagcac acagcaggca gctcacgaat tctcttttcc 63300
tgttctccct tgggagctgg gaggatctta atcaggcaat aagagatggc actgagcagc 63360
cagctaattt tttaaatcac tttattgttt aaccatatga ctcacccact taaaaaaggg 63420
tacagttcag tgggttttag tgtattcaca gatgtgtgca accctcacca cagttaattt 63480
tagaacattt tcctgcccct aaaagaaact ctgcatgaag ccagctgttt ttaaattagc 63540
aaagttattt tgcatccttt aaatatatgt tcatggtaca aaattcaaaa gatacagaag 63600
agtctgcagt ccaaagagac tccgccccca tgacgccaag caggactccc tgggaggcat 63660
ggcctcctgc agtgtgtttc ttctatgtcc ccccaggggt catctgtaca tatgcaagca 63720
tacaagagcg tggactttgt tttccaagcc agaagataat tgtagattta tgtgcagttg 63780
tgagaaagag cacagaccca tttatcctct gcctggtttc cccagtgct gcctgccatc 63840
ttgcatgact tccattccta tcataagcaa gacactgata acgattcttt caccttattc 63900
agattgacat aagtgttttt tgtttgttct tgagacaaac ttcctctgtc acccagtggg 63960
agtgcagtgg cacaatcaca gctcactgca gcctcaaact cctgggctca agcgattctc 64020
ctgcctcagt cccctcaagt agctcagatg gcaggtgtgc accatcatgc caggctaatt 64080
tttaaatttt ttgtggaggt gaggcctcac taaatttcct gggctagtct tgaactcctg 64140
agctaaagtg atcctcctgc ctcagcctcc caaagtggta ggattacagg catgagccac 64200
tgcgcctggg ctgacatatg tgttttcgta agcccgaaag atagcatctg aagagtcaac 64260
attgagcctt gccttttgct gctaatgatg tataaaagct gctgttctga gcatttcgga 64320
ggctcccagc tgccgtgtgc accctgccta gagctctacc gtaacccatc tccgggagga 64380
ggtgctattg ttttcctcat tttgcaacaa ggaggctgaa gaactgagca tgaaccactg 64440
gcctgggtcg ttcggttggt aggcagtggg gccaggccat ccaactcaca accaccttct 64500
actctgcttc ccccgcaccc tgaagtttgt tctgttttga ggacacagcc gtcacattct 64560
tggtggctga acagcactcc ttgtcaggtg tggctgggcc cccactggag ggcatcatgg 64620
tcctctctcc tgctgcggtt gaaccttggc tgtttcaacc actcctgcca agtggccctc 64680
tgaaagggac agtccatctt ttctcagcag agggccacac tggcaaaacg gtccctggca 64740
ccctttctct ccacctgtct aatatagagt aaaaatggta tcatgttaag atcttcattt 64800
atatttattt tatcatgaat gatgtaagca tcattttgtg tgtttaagaa cctttgggcc 64860
cagcgtgatg gcttgcagct gtaatctcag cactttagga ggctgagatg agcggatcac 64920
ttgaggccgg gagtttgaga ccagcctggc caacatggag aaaccccgtc tctagtaaaa 64980
atttaaaaat tagccgggta tggtgatccc agctacttgg gagtctgaag catgagaatt 65040
gcttgaacat gggaggcgga ggttgcagtg agccgagatc gcgccattgc actccagcct 65100
gggcgacaga gcgagactct gtctcacaaa aaaaaaaaaa aaagaaaaga aaagaaatta 65160
tcaatctcct cttttatggc atatatatat atatatatat atatatatat ttatttccct 65220
ttcttggtta tgttcataaa ggcctcccct gctctgatca taaaaaacaa cttattttca 65280
cactctctct cttttttttt tgagacagag ttttgctcct gttgcccagg ctggagtgca 65340
gtggcgcaat ctcagctcac tgtaacctcc gcctcccggg ttggagtgat tctcctgcct 65400
taccttcccg agtagctggg attataggca tgcaccacca tgcctggcta attttgtact 65460
tttagtagag acggggtttt ctccatgttg gtcaggctgg tctcgaactc gcgacctcag 65520
gtgatccacc cacctcggcc tcccaaagtg ctgggattac agacgtgagc caccatgccc 65580
agcccacact ctctttctta acgtcctcct cctttcgttt tacgttcaca tctttaattc 65640
ttctgggatg taattagatt tgatgagcaa ggtgggcatc cagcttgttt cttggctgat 65700
ggcttatggg tggcgtgaat tagtcggggt ctatcaggag gcagaaactc tatgagaatt 65760
tgaacagaga aagttccgtc tacaggctta ttaccaggga ctggaatagc agaaattgaa 65820
cagtgagatg tacagagaac tctaagaatg caggaatagg ccaggcatgg tggctcacac 65880
ctgtcatccc agcactttgg gagaccaagg cgggtggatc acctgaggtc aggagttcga 65940
gaccagcctg gccaacatag tgaaacccca tctctactaa aaatacaaaa aaattagctg 66000
ggtgtggtgg cgcatgcctg taatcccagc ttctcgggag tctgaggctg gagaatcact 66060
tgaacctggg aggcagaggt tgtagtgagc cgagatcatg ccattgtact ccagcctggg 66120

caacaagagc gagactcagt caaaacaaca acaacgcagg aatagcagat gagccgaggt 66180
ggggcctccc cagcccccac cccccacccc gcaccctggg ccgagatcca gtcctctttg 66240
aatagggcct gggcgtggtt cacgggacat ctgagacatt gccgaggcgc tgcactggtg 66300
gatcttgcca gaagtctgcc cagtgcagat ttgggcagaa tctcaaactg ccttgggatg 66360
taggagagaa accaggcctg gtcaagttca tgggaagagg tggaaacaga ccccataggc 66420
tggggcttgg gcagctgtag gaagccctct ctgctgcctc cctgcctgct ctctgctttg 66480
aagcatcttc cccagtgccc ccagtctcat gccctctcaa cgttggggtc aaatcctgag 66540
gaatacccag actggctctc tgggccaaag aggaccctct ccagaaagag cagggcccag 66600
tgcggcttcc taaagggcag gggaagggcc tggccactcc ccagaggcta ctcaccagcc 66660
atcaggatag ccccaggaag caggccttct cgagcccatt ttattacttt attttattat 66720
tttatttaat tttaaattta ttttttgaga cagagtctca ctctgttgcc caggctggag 66780
tgcagtggtg cgatctcaac ccactgcagc ctctgcctcc agggttcaag ggattctccc 66840
acctcagcct cccaagtagc tgggattaca ggtgcccgcc accacacccg gctaattttc 66900
atatttttag tagagacgag gtttcaccat gttggccagg ctggtctcga actcctgacc 66960
tcaagtgatc cgcccgcctc ggcctcccaa agtgctaggt caagcccatt ttaaagttga 67020
agaaactgag gctgaggtaa attccctccc cagggatcct gctgcagcca gaaggtggta 67080
aaacaggact tcacccgggt ctgtctggcg tgaaaggcag tgttcttgta ccaccctagg 67140
gggcctgaga gaactgagtc cctcgggcat aactgacagt tctgttccca ttattccgca 67200
ggggctcgga tctggctgta tgctttccag gatggccttg gagacccaca taagccctac 67260
accctttggg aagctgcatg ttgggttggg gtgccgtcag tggcacttgt ggaaggtgca 67320
gacctgtgtg ggtgtgtggg cccagggccc ctggtccctt cctccctttg tagggctggt 67380
tgtgtgctgc ctggacctgg ggggcacgtt cacgtggtga atttgtctat ttactatccc 67440
cgctttgggg ctggtgccag cacaggccct tgtgaagggg gtgcctttgt ctggagtggg 67500
actgtggccc ctccctcagc gtggtgactt ctgtgtcagg gcttcagcag ggacgcagag 67560
cccctgagtg ttcggaacaa gggcgtcatt gcaggagtta gactgtgtgt gatggaggga 67620
ggagggggcag gaggaaaggt cagaaggaga gttcctggga aggtccctga ggagcctggt 67680
gaggtgctaa ctggtgtgga ggacactcag ggcctgtggg gacatctcct actgctgggg 67740
gccagccaca aagggaactg gccgaagtcc tgtccccgcc ttcacagccc agcatctggt 67800
cacaaggcag gtacttggaa gggcgcgggc acctgggcca aaagtgcctg ggttcccttt 67860
gcctttcact gagatgacct tcggggcagg tggctgctgc ctcccctcct gtccccaggt 67920
tttgccaact ggccagagga aggggtcctg ggaagcaggg gggccagaag ccctctctgc 67980
aaggaaagcc cgaggggtgt gggaggaagg aaggaatgcc caggctggcg aggctctaag 68040
tcaccctggc ttggctctcc tcagatcctg aacccgccgc cctccccggc cacggacccc 68100
tccctgtaca acatggacat gttctactct tcaaacattc cggccactgc gagaccgtac 68160
aggtaggaca tcccctgcag ccctccatgg ccattgggtt cccgccagcc cgtggtggag 68220
gggcctaatc cccatgccac tgatgagggg aggtattctg ggtgctagtg ggcaggtgcc 68280
gggcccagcc ctgcctccct ctgctctgcc aaccacacta ggctgcctcc ccagacaagc 68340
tcagcgggca ctgcatgttg ggttcagaaa tcagcagaac tccacgttct gagctgctct 68400
tcaagttgct cctatggggg ttacttttaa gctgggaaat ggctgtggcg tcgagggggcc 68460
gggggcttgg gctccaaact ctgactgtgt gtttgagtcc ggctgtggaa acctagccat 68520
tgagatgccc cctcttggtg gctctgtcct cttaggatgg gacaagtctg tgaaggctgc 68580
tgcagcaccc accgtagacc cctaatcgtg tgacgtcacc aggatggtcc gggctgctca 68640
cttgccacag tggcctgttt gagcccggga agccaacggg gctgctcagc tggacaccag 68700
ccccccgagc tgcccatgtt ggggtcacag gccccacctc cctggttggg gaggggcaac 68760
tgagagtgtg gagaggtggg acccaggtgt gctggtctcc gcagggggctg gatcagagcc 68820
tgggatgggc agggtgagcc tcctgacctt taacccagtg gtgtcaggca acgtggccca 68880
cccgccagcc gcaccaggcc ccaccccccgc aggtgaaggg gtgggatagg ctgggcctgg 68940
gccaggacac ctctggacca cgcattcctc attgcttggg tccctggagc agcagggcct 69000
cccgagtgtg gtgccgcctg ccacctagtg gccatttcca cgaactccca ggcctggctg 69060
gggagccgga actgcagcct ccatttccac cccactccgg gtcgggccac ctccctgatg 69120
cctcagtatt atatcaaact gtcacagtct gtcccacagc cttacagacc actgtctcca 69180
gaatggtcac atccacactg ggcagcccag tctcgctagt tcctcgtccc acctcctgcc 69240
tttgctcatg cccgtcctgc tctgggccca ccgcggacac atcttccccc cgcccgccgt 69300
ctgacctcac agcagctggg ccccaagagg agtatcctgt cctgctgcac ttttctcaac 69360
acccggtgtt ggctgcacct tcccacccat tgcaggcccc tctgtgacag gacgggggct 69420
cctaaacaca ccacagttcc gagtctgaac tcacacagtg ggatgcggcg tttctgggcc 69480
acagttgggt gcaggtagcc tctgggagga tgggaggtca ggagccatct tgcgagtcag 69540

```
gttgcttgaa ctcaggatgg aagtgttccg ggcccattgg ttgctgtatt agcctgttct     69600
cacgctgcta ataaagacat acccaagact gggtaattgt aaaggaaaga ggtttaacgg     69660
actcacagtt ccacctgcct ggggtggcct cacaatcatg gtagaagaca aggaggagca     69720
agtcacatct tacatggctt cagggaacag acagcatgag aaccaagcga aaggggtttc     69780
cccttgtaaa accatcaagt ctagtgagat ttattcacta ccacgagaac agtatggggg     69840
gaaccacccc catgattcaa tcatctccca ctgggtccct cccacagcac gtgggaatta     69900
tgggagtaca attcaagatg agatttgggt ggggacacag ccaaacccta tcggttgcca     69960
acatttacag taacagtgtt aggtgaacag ttgtccagtc tcctgttttg tcggacactg     70020
tttctagcac cttccaggca gaatctcatg tatccttcac tttcgaaatg ggtactattt     70080
catccccact tttatcaatg agaaactaaa gctcgaagag gtcaagtaag ttcctggcca     70140
aggtcagcta gcaggctcta gaggcctcgt tctccttaga ggcagccttg ccagggccca     70200
ggcttggcag gctgcagggc aggtgcgggc atgcccatgg tagaggtggg accattgagg     70260
ctcagagagg gtaagtgatg agccctggcg acacagcggg gtgggtccag agtccggcct     70320
gcatcttctg gagctggcca gtggacaggc ctttcccgtt cacagccccg gggctgctgt     70380
gcccaccagg gcggatgtgc ctaccgaatc ccactcctct gtgtgtgtcc ctttcaggcc     70440
ctacatcatt cgaggaatgg cgcccccgac gacgccctgc agcaccgacg tgtgtgacag     70500
cgactacagc gccagccgct ggaaggccag caagtactac ctggatttga actcggactc     70560
agacccctat ccaccccccac ccacgcccca cagccagtac ctgtcggcgg aggacagctg     70620
cccgccctcg cccgccaccg agaggagcta cttccatctc ttcccgcccc ctccgtcccc     70680
ctgcacggac tcatcctgac ctcggccggg ccactctggc ttctctgtgc ccctgtaaat     70740
agttttaaat atgaacaaag aaaaaaatat attttatgat ttaaaaaata aatataattg     70800
ggattttaaa aacatgagaa atgtgaactg tgatggggtg ggcagggctg ggagaacttt     70860
gtacagtgga gaaatattta taaacttaat tttgtaaaac agaactgcca ttcttttgtg     70920
ccctgtgtgc atttgagttg tgtgtccccg tggagggaat gccgaccccc ggaccaccat     70980
gagagtcctc ctgcacccgg gcgtccctct gtccggctcc tgcagggaag ggctggggcc     71040
ttgggcagag gtggatatct cccctgggat gcatccctga gctgcaggcc gggccggctt     71100
tatgtgcgtg tggcctgtgc cgtcagaaag ggccctgggc ttcatcacgc tgttgctgtt     71160
cgtcttcctc agattcttag tctttttttt tttttttttt ttttgagacg gagtctttct     71220
ctgtcatcca ggctggagtg cagtggtaca atctcagctc actgcaagct ccgactccca     71280
ggttcaagtg agtctcctgc ctcagcctcc cgagtagctg ggactacagg tgcgcgccac     71340
cacacccgcc cagctaattt ttgtattttt agtagagatg gggtttcacc atgttggcca     71400
ggatgatctc gatctcttga cctcgtgatc cgcccacctc ggcctcccaa agtgctggga     71460
ttataggcat gagccactgt acccagctga ctcttagtca cttttaagaa ggggactgtg     71520
ccttcatttt tcactgggcc ctgcagaata tatgcctggg ctctgggctc ttctgaacct     71580
gtgttggctt ccatctgacc tctctgtgcc agcccaaggc tgctgctctt cctgagggca     71640
aggagcccca tgactgcgtg ttgactcgct ggatggggct gctgagccca ctctgccaca     71700
ccacgtgccc ctggcaggga gggaatccct gggtcctcac aggaacagtc agcaagccac     71760
acctgacgcc tgctgtgggc ccatccctgc ggtgctggag aagacagaca aggcctggtc     71820
actgcctctg cagggtcccc agtccgtgga aggagacagt aatctaggca ttttcggtgg     71880
ggaagctgag ctgttctcgt gtcctgaagg ccaggcggga acagccgtct tcagagggaa     71940
gggagaaaat gcacatcgca tcagtggaga agggcctgac ttccctcagc atggtggagg     72000
gaggtcagaa aacagtcaag cttgagtatt ctatagtgtc acctaaata                 72049
```

<210> 10
<211> 8705
<212> DNA
<213> Homo sapiens

<400> 10

```
ggactcaggg gcagcaggga ggtacaccca tggttagtgg gcggaccata gggggtaatg     60
agagggtgaa tcgatggaac ctgggggaca caatcgaagt ggttccagag tcgggctgta     120
ctaattaaag agacggggca gtggacaggc attttcagtt gactgcccag ggagtgttct     180
gcccaacagg gaggatatgc gtacagaatc atactcgatc agcatgagtc caattcagac     240
cgtacatcag tggagatatg ggtcccccga tgactccgtg gaacactgat gtttgtgaca     300
ggggagtaca gcaccagcca tcagcaggcc agtaaatcat accggcctgc gaaattggac     360
```

```
tcagacccgg atccaccctg accgacgtcc caagccccca ccccccaccc cccaccatgg     420
gccgagatcc agtcctcttt gaatagggcc tggccgtggt tcacgggaca tctgagacat     480
tgccgaggcg ctgcattggt ggatcttgcc agaagtttgc ccagtgcaga tttgggcaga     540
atctcaaact gccttgggat gtaggagaga aaccaggcct ggtcaagttc atgggaagag     600
gtggaaacag accccatagg ctggggcttg ggcagctgta ggaagccctc tctgctgcct     660
ccctgcctgc tctctgcttt gaagcatctt ccccagtgcc cccagtctca tgccctctca     720
acgttggggt caaatcctga ggaataccca gactggctct ctgggccaaa gaggaccctc     780
tccagaaaga gcagggccca gtgcggcttc ctaaagggca ggggaagggc ctggccactc     840
cccagaggct actcaccagc catcaggata gccccaggaa gcaggccttc tcgagcccat     900
tttattactt tattttatta ttttatttaa ttttaaattt attttttgag acagagtctc     960
actctgttgc ccaggctgga gtgcagtggt gcgatctcaa cccactgcag cctctgcctc    1020
cagggttcaa gggattctcc cacctcagcc tcccaagtag ctgggattac aggtgcccgc    1080
caccacaccc ggctaatttt catattttta gtagagatga ggtttcacca tgttggccag    1140
gctggtctcg aactcctgac ctcaagtgat ccgcccgcct cggcctccca aagtgctagg    1200
tcaagcccat tttaaagttg aagaaactga ggctgaggta aattccctcc ccagggatcc    1260
tgctgcagcc agaaggtggt aaaacaggac ttcacccggg tctgtctggc gtgaaaggca    1320
gtgttcttgt accaccctag gggcctgag agaactgagt ccctcgggca taactgacag    1380
ttctgttccc attattccgc aggggctcgg atctggctgt atgctttcca ggatggcctt    1440
ggagacccac ataagcccta caccctttgg gaagctgcat gttgggttgg ggtgccgtca    1500
gtggcacttg tggaaggtgc agacctgtgt gggtgtgtgg gcccagggcc cctggtccct    1560
tcctcccttt gtagggctgg ttgtgtgctg cctggacctg gggggcacgt tcacgtggtg    1620
aatttgtcta tttactatcc ccgctttggg gctggtgcca gcacaggccc ttgtgaaggg    1680
ggtgcctttg tctggagtgg gactgtggcc cctccctcag cgtggtgact tctgtgtcag    1740
ggcttcagca gggacgcaga gcccctgagt gttcggaaca agggcgtcat tgcaggagtt    1800
agactgtgtg tgatggaggg aggaggggca ggaggaaagg tcagaaggag agttcctggg    1860
aaggtccctg aggagcctgg tgaggtgcta actggtgtgg aggacactca gggcctgtgg    1920
ggacatctcc tactgctggg ggccagccac aaagggaact ggccgaagtc ctgtccccgc    1980
cttcacagcc cagcatctgg tcacaaggca ggtacttgga agggcgcggg cacctgggcc    2040
aaaagtgcct gggttccctt tgcctttcac tgagatgacc ttcggggcag gtggctgctg    2100
cctcccctcc tgtccccagg ttttgccaac tggccagagg aaggggtcct gggaagcagg    2160
ggggccagaa gccctctctg caaggaaagc ccgaggggtg tgggaggaag gaaggaatgc    2220
ccaggctggc gaggctctaa gtcaccctgg cttggctctc ctcagatcct gaacccgccg    2280
ccctccccgg ccacggaccc ctccctgtac aacatggaca tgttctactc ttcaaacatt    2340
ccggccactg cgagaccgta caggtaggac atcccctgca gccctccatg gccattgggt    2400
tcccgccagc ccgtggtgga ggggcctaat ccccatgcca ctgatgaggg gaggtattct    2460
gggtgctaat gggcaggtgc cgggcccagc cctgcctccc tctgctctgc caaccacact    2520
aggctgcctc cccagacaag ctcagcgggc actgcatgtt gggttcagaa atcagcagaa    2580
ctccacgttc tgagctgctc ttcaagttgc tcctatgggg gttactttta agctgggaaa    2640
tggctgtggc gtcgaggggc cgggggcttg ggctccagag tctgactgtg tgtttgagtc    2700
cggctgtgga aacctagcca ttgagatgcc ccctcttggt ggctctgtcc tcttaggatg    2760
ggacaagtct gtgaaggctg ctgcagcacc caccgtagac ccctaatcgt gtgacgtcac    2820
caggatggtc cgggctgctc acttgccaca gtggcctgtt tgagcccggg aagccaacgg    2880
ggctgctcag ctggacacca gcccccgag ctgcccatgt tggggtcaca ggccccacct    2940
ccctggttgg ggagggggcaa ctgagagtgt ggagaggtgg gacccaggtg tgctggtctc    3000
cgcaggggct ggatcagagc ctgggatggg cagggtgagc ctcctgacct ttaacccagt    3060
ggtgtcaggc aacgtggccc acccgccagc cgcaccaggc cccacccccg caggtgaagg    3120
ggtgggatag gctgggcctg ggccaggaca cctctggacc acgcattcct cattgcttgg    3180
gtccctggag cagcagggcc tcccgagtgt ggtgccgcct gccacctagt ggccatttcc    3240
acgaactccc aggcctggct ggggagccgg aactgcagcc tccatttcca ccccactccg    3300
ggtcgggcca cctccctgat gcctcagtat tatatcaaac tgtcacagtc tgtcccacag    3360
ccttacagac cactgtctcc agaatggtca catccacact gggcagccca gtctcgctag    3420
ttcctcgtcc cacctcctgc ctttgctcat gcccgtcctg ctctgggccc accgcggaca    3480
catcttcccc ccgcccgccg tctgacctca cagcagctgg gccccaagag gagtatcctg    3540
tcctgctgca cttttctcaa cacccggtgt tggctgcacc ttcccaccca ttgcaggccc    3600
ctctgtgaca ggacgggggc tcctaaacac accacagttc cgagtctgaa ctcacacagt    3660
gggatgcggc gtttctgggc cacagttggg tgcaggtagc ctctgggagg atgggaggtc    3720
aggagccatc ttgcgagtca ggttgcttga actcaggatg gaagtgttcc gggcccattg    3780
```

```
gttgctgtat tagcctgttc tcacgctgct aataaagaca tacccaagac tgggtaattg 3840
taaaggaaag aggtttaacg gactcacagt tccacctgcc tggggtggcc tcacaatcat 3900
ggtagaagac aaggaggagc aagtcacatc ttacatggct tcagggaaca gacagcatga 3960
gaaccaagcg aaaggggttt ccccttgtaa aaccatcaag tctagtgaga tttattcact 4020
accacgagaa cagtatgggg ggaaccaccc ccatgattca atcatctccc actgggtccc 4080
tcccacagca cgtgggaatt atgggagtac aattcaagat gagatttggg tggggacaca 4140
gccaaaccct atcggttgcc aacatttaca gtaacagtgt taggtgaaca gttgtccagt 4200
ctcctgtttt gtcggacact gtttctagca ccttccaggc agaatctcat gtatccttca 4260
ctttcgaaat gggtactatt tcatccccac ttttatcaat gagaaactaa agctcgaaga 4320
ggtcaagtaa gttcctggcc aaggtcagct agcaggctct agaggcctcg ttctccttag 4380
aggcagcctt gccagggccc aggcttggca ggctgcaggg caggtgcggg catgcccatg 4440
gtagaggtgg gaccattgag gctcagagag ggtaagtgat gagccctggc gacacagcgg 4500
ggtgggtcca gagtccggcc tgcatcttct ggagctggcc agtggacagg cctttcccgt 4560
tcacagcccc ggggctgctg tgcccaccag ggcggatgtg cctaccgaat cccactcctc 4620
tgtgtgtgtc cctttcaggc cctacatcat tcgaggaatg gcgcccccga cgacgccctg 4680
cagcaccgac gtgtgtgaca gcgactacag cgccagccgc tggaaggcca gcaagtacta 4740
cctggatttg aactcggact cagaccccta tccaccccca cccacgcccc acagccagta 4800
cctgtcggcg gaggacagct gcccgccctc gcccgccacc gagaggagct acttccatct 4860
cttcccgccc cctccgtccc cctgcacgga ctcatcctga cctcggccgg gccactctgg 4920
cttctctgtg cccctgtaaa tagttttaaa tatgaacaaa gaaaaaaata tattttatga 4980
tttaaaaaat aaatataatt gggattttaa aaacatgaga aatgtgaact gtgatggggt 5040
gggcagggct gggagaactt tgtacagtgg agaaatattt ataaacttaa ttttgtaaaa 5100
cagaactgcc attctttcgt gccctgtgtg catttgagtt gtgtgtcccc gtggagggaa 5160
tgccgacccc cggaccacca tgagagtcct cctgcacccg ggcgtccctc tgtccggctc 5220
ctgcagggaa gggctggggc cttgggcaga ggtggatatc tcccctggga tgcatccctg 5280
agctgcaggc cgggccggct ttatgtgcgt gtggcctgtg ccgtcagaaa gggccctggg 5340
cttcatcacg ctgttgctgt tcgtcttcct cagattctta gtcttttttt tttttttttt 5400
tttttgaga cggagtcttt ctctgtcatc caggctggag tgcagtggta caatctcagc 5460
tcactgcaag ctccgactcc caggttcaag tgagtctcct gcctcagcct cccgagtagc 5520
tgggactaca ggtgcgcgcc accacacccg cccagctaat tttgtattt ttagtagaga 5580
tggggtttca ccatgttggc caggatgatc tcgatctctt gacctcgtga tccgcccacc 5640
tcggcctccc aaagtgctgg gattataggc atgagccact gtacccagct gactcttagt 5700
cacttttaag aaggggactg tgccttcatt tttcactggg ccctgcagaa tatatgcctg 5760
ggctctgggc tcttctgaac ctgtgttggc ttccatctga cctctctgtg ccagcccaag 5820
gctgctgctc ttcctgaggg caaggagccc catgactgcg tgttgactcg ctggatgggg 5880
ctgctgagcc cactctgcca caccacgtgc ccctggcagg gagggaatcc ctgggtcctc 5940
acaggaacag tcagcaagcc acacctgacg cctgctgtgg gcccatccct gcggtgctgg 6000
agaagacaga caaggcctgg tcactgcctc tgcagggtcc ccagtccgtg gaaggagaca 6060
gtaatctagg cattttcggt ggggaagctg agctgttctc gtgtcctgaa ggccaggcgg 6120
gaacagccgt cttcagaggg aagggagaaa atgcacatcg catcagtgga gaagggcctg 6180
acttccctca gcatggtgga gggaggtcag aaaacagtca agcttgttgc tgggtgacag 6240
tgcatttaat aatcaaaata taggctgggt acggtggctc atgcctgtaa tcccagcact 6300
ttgggaggct gaggcaggtg gatcacttga ggccaggagt ttgagaccgg cctggccaac 6360
atggcaaaac ctcaactact aaaatacaaa aactagccgg gcgtggtggt gcacgcctgt 6420
aatcccagct acttgggagg ctgaggcagg agaattgctt gaacctggga ggcggaggct 6480
gcagtgagcc gagattgtgc cactgcactc cagcctgggc aacagagcaa gactctgtct 6540
caaaaaaaaa aaaaaaaaaa gcaatacaaa atacaaatat cactttcact aaaagaaggg 6600
atggaagacc caaaacaaac agaaaacaac aaaatggcag gagtaagtcc ccacttatca 6660
ataataacat tgactgtaaa taggctaagc tctgcaatca aaagagtggg ccaggagcgg 6720
tggctcacgc ctgtaattcc aacgctttgg gaggctgagg cggatggatc atttgatgtc 6780
acgagtttta agaccagcct ggccaacaag gtgaaacccc atctgtacta aaaatacaaa 6840
aattagccag gcggtagtgg cacgcacctg taatcccagc tacttgtgag gctgaggcag 6900
gagaatcact ggaggctggg aagcggaggt tgctgtgagc caagatggag ccactgcact 6960
cccacctggg cgacagagtg agatcctgtc ttaagaaaaa aaagagtgga tgaatggatc 7020
aaaaaacaag acccaaccat ctcttgcata caagaaacac actttaccta taaaaacaca 7080
ctaggccagg tgtggtggct cacacctgta atcccagccc tttgggaggc ctgactggca 7140
gatcacctga ggccaggagt ttcagaccag cttgaccgac atggcaaaac cccatctctc 7200
```

ctaaaaatac aaaaaaacaa aaaaaagaaa aaggctggaa gtagtgatgt gtgcctgtag 7260
ccccagctac ttgggaggct gaggcaggag aattgcttga atccgggaag tggaggttgc 7320
agtgagccag gatggtgcca ctgcactcca gcctgggtga cagagcgaga ccctgtcata 7380
aaaaaaaaaa gaaaagaaaa gaaaaacgag aaaaacaaac acaaaattag tagaagaaaa 7440
gaaataataa agatcagaac aggccaggct catgggcaca gtggctcaac tcctacctgc 7500
tcaggagttt gagaccagtc tggccaacat ggcaaaaccc catctctcct aaaaatatga 7560
aaaaaaaaaa ataggctgga tgtggtgatg tgtgtgtgcc tgtagcccca gctacttggg 7620
aggctgaggt gggagaatca cttgagccca ggaagtggag gctgcagcga gtcatgaatg 7680
caccctgcac tctagctggg taactggagt gagattctgt ctcaaaaaag caaagaccag 7740
agcagaaata aatgaaatgg aaatgaagga aacaatgcaa aatgatacaa aaagtttttt 7800
cgaaaagata aacaaaatca acaaaccttt agccagatta agaaaaaaag agagaagacc 7860
caaataaata aaatccgaga ttaaaaagga gacattacca ctgataccac agaaattcaa 7920
aggatcatta gaggcaacta tgtgcaacta tatgctaatg aactggaaaa cctagaagaa 7980
ctgggtaaat ttctagacac atacaaccta tcaagattga accatgaaga aatccaaaac 8040
ctgaacaggc cgggcacggt ggcttacgcc tgtaatccca gcactttgga aggcctgaga 8100
tcaggagttc gagaccagcc tggccaacat ggtgaaaccc catctctact gaaaaaatat 8160
aaaaattagc cgggcgtggt ggcgggtgcc tctaatgtca gccactcggg aggctgaggc 8220
aggaaaatca cttgaacctg ggaggcatag gttgcagcga gccgaggttg caccactgca 8280
ctccagcctt ggcgacagag ccagactcca tctcaaaaaa attaaaataa caaaaacctg 8340
aacagaccaa taacaagtaa tgcgatgaaa actgtaataa aatgtttccc aacaaagaaa 8400
gcccaggaac aaatggcttc actgctgaat tttaccaaac attttttttt ttttgagacg 8460
gagtctcgct ctgtcgccca ggctggagtg cagtggtgta acctcggttc gctggtaact 8520
tatgcctctc aggctgcaag tgattttcct gcttcaggcc ccccgagtgg ctggaaatta 8580
gatggtactt gtcaaacaag gcctggctaa atttctatat ttccttcaag tagaagatgt 8640
gcttccaaca aaggttgggt tacggctggc ttctgaaaat cttggatttc aaggctcccc 8700
aaaag 8705

<210> 11
<211> 66933
<212> DNA
<213> Homo sapiens

<400> 11
tataatcaag cgcgttccgt ccagtccggt gggaagattt tcgatatgct tcgtgatctg 60
ctcaagaacg ttgatcttaa agggttcgag cctgatgtac gtattttgct taccaaatac 120
agcaatagta atggctctca gtccccgtgg atggaggagc aaattcggga tgcctgggga 180
agcatggttc taaaaaatgt tgtacgtgaa acggatgaag ttggtaaagg tcagatccgg 240
atgagaactg tttttgaaca ggccattgat caacgctctt caactggtgc ctggagaaat 300
gctctttcta tttgggaacc tgtctgcaat gaaattttcg atcgtctgat taaaccacgc 360
tgggagatta gataatgaag cgtgcgcctg ttattccaaa acatacgctc aatactcaac 420
cggttgaaga tacttcgtta tcgacaccag ctgccccgat ggtggattcg ttaattgcgc 480
gcgtaggagt aatggctcgc ggtaatgcca ttactttgcc tgtatgtggt cgggatgtga 540
agtttactct tgaagtgctc cggggtgata gtgttgagaa gacctctcgg gtatggtcag 600
gtaatgaacg tgaccaggag ctgcttactg aggacgcact ggatgatctc atcccttctt 660
ttctactgac tggtcaacag acaccggcgt tcggtcgaag agtatctggt gtcatagaaa 720
ttgccgatgg gagtcgccgt cgtaaagctg ctgcacttac cgaaagtgat tatcgtgttc 780
tggttggcga gctggatgat gagcagatgg ctgcattatc cagattgggt aacgattatc 840
gcccaacaag tgcttatgaa cgtggtcagc gttatgcaag ccgattgcag aatgaatttg 900
ctggaaatat ttctgcgctg gctgatgcgg aaaatatttc acgtaagatt attacccgct 960
gtatcaacac cgccaaattg cctaaatcag ttgttgctct tttttctcac cccggtgaac 1020
tatctgcccg gtcaggtgat gcacttcaaa aagcctttac agataaagag gaattactta 1080
agcagcaggc atctaacctt catgagcaga aaaaagctgg ggtgatattt gaagctgaag 1140
aagttatcac tcttttaact tctgtgctta aaacgtcatc tgcatcaaga actagtttaa 1200
gctcacgaca tcagtttgct cctggagcga cagtattgta taagggcgat aaaatggtgc 1260
ttaacctgga caggtctcgt gttccaactg agtgtataga gaaaattgag gccattctta 1320

```
aggaacttga aaagccagca ccctgatgcg accacgtttt agtctacgtt tatctgtctt    1380
tacttaatgt cctttgttac aggccagaaa gcataactgg cctgaatatt ctctctgggc    1440
ccactgttcc acttgtatcg tcggtctgat aatcagactg ggaccacggt cccactcgta    1500
tcgtcggtct gattattagt ctgggaccac ggtcccactc gtatcgtcgg tctgattatt    1560
agtctgggac cacggtccca ctcgtatcgt cggtctgata atcagactgg gaccacggtc    1620
ccactcgtat cgtcggtctg attattagtc tgggaccatg gtcccactcg tatcgtcggt    1680
ctgattatta gtctgggacc acggtcccac tcgtatcgtc ggtctgatta ttagtctgga    1740
accacggtcc cactcgtatc gtcggtctga ttattagtct gggaccacgg tcccactcgt    1800
atcgtcggtc tgattattag tctgggacca cgatcccact cgtgttgtcg gtctgattat    1860
cggtctggga ccacggtccc acttgtattg tcgatcagac tatcagcgtg agactacgat    1920
tccatcaatg cctgtcaagg gcaagtattg acatgtcgtc gtaacctgta gaacggagta    1980
acctcggtgt gcggttgtat gcctgctgtg gattgctgct gtgtcctgct tatccacaac    2040
attttgcgca cggttatgtg gacaaaatac ctggttaccc aggccgtgcc ggcacgttaa    2100
ccgggctgca tccgatgcaa gtgtgtcgct gtcgacgagc tcgcgagctc ggacatgagg    2160
ttgccccgta ttcagtgtcg ctgatttgta ttgtctgaag ttgtttttac gttaagttga    2220
tgcagatcaa ttaatacgat acctgcgtca taattgatta tttgacgtgg tttgatggcc    2280
tccacgcacg ttgtgatatg tagatgataa tcattatcac tttacgggtc ctttccggtg    2340
atccgacagg ttacggggcg gcgacctcgc gggttttcgc tatttatgaa aattttccgg    2400
tttaaggcgt ttccgttctt cttcgtcata acttaatgtt tttatttaaa ataccctctg    2460
aaaagaaagg aaacgacagg tgctgaaagc gagctttttg gcctctgtcg tttcctttct    2520
ctgtttttgt ccgtggaatg aacaatggaa gtccgagctc atcgctaata acttcgtata    2580
gcatacatta tacgaagtta tattcgatgc ggccgcaagg ggttcgcgtc agcgggtgtt    2640
ggcgggtgtc ggggctggct taactatgcg gcatcagagc agattgtact gagagtgcac    2700
catatgcggt gtgaaatacc gcacagatgc gtaaggagaa aataccgcat caggcgccat    2760
tcgccattca ggctgcgcaa ctgttgggaa gggcgatcgg tgcgggcctc ttcgctatta    2820
cgccagctgg cgaaaggggg atgtgctgca aggcgattaa gttgggtaac gccagggttt    2880
tcccagtcac gacgttgtaa aacgacggcc agtgaattgt aatacgactc actatagggc    2940
gaattcgagc tcggtacccg gggatcctct agagtcgacc tgcaggcatg caagcttctc    3000
ttgtgccggt tgtacgctgt caggtcacac tggtgagtta ggcagggcac agatgcccag    3060
agcagaggga actttccttg gggattcaac acgtgcaagt cttaggggct ggcaaatcct    3120
gccctcagct agagaggggg cttttatttg agaccagaat cacctgagca tcctcctgtc    3180
cccagctgtg tccagcctgt ctgcagggac atcctgagag gaccaggctc tcccctcatc    3240
cacctgccta agtgccactc tgaaccctgt ccacctgtgc cgtggagggg cgtgacctca    3300
agctgctcag ccagcagcag gcttggccct gggggcagc agagacccag gtggctgtgg    3360
ggtgggtgct tcgtggcgtg gttctgaaac ttcgttggaa gtgtgtggac agtgccttgc    3420
ctgttctctg tgggacccta tttagaaacg aggtctgagt tactgggggt catcactgtg    3480
ttctgatggc ccagctgtgt ggaggccgcg gtgcagcccc atccaaggag ccagggccct    3540
gggtctagcc gtgaccagaa tgcatgcccc ggaggtgttt ctcatctcgc acctgtgttg    3600
cctggtgtgt caagtggtcg tgaaactctg tgttagctct tggtgttcct gaaagtgccc    3660
ccgggtctca ggcctcagaa ccagggtttc ccttcatctc ggtggcctgg gagcatctgg    3720
gcagttgagc aaagagggcg attcacttga aggatgtgtc tggccctgcc taggagcccc    3780
ccggcacggt gctggggcct gaagctgccc tcgggtggtg gagaggaggg agcgatgaag    3840
tggcgtcgag ctgggcagga agggtgagcc cctgcaaggt gggcatgctg ggacgctga    3900
gcagcatggc cagcagctgg gtctgcagcc tggtacccgg cgggacttgt ggttggggct    3960
ggtttgtggc caggagaggg gctggcagga gacaaggggg actgtgaggc agctcccacc    4020
cagcagctga agcccaatgg cctggctgtg tggctctcag ctgcgtgcat aacctctcag    4080
tgcttcagtt ctctcatttg taaaatgagg aaacaaacag tgccagcctc ccagaggtgt    4140
catgaggatg aacgagtgac catgtagcat gggctgggtg cgtgtcacct aacatcacca    4200
gcctttgcaa ggagagccct gggggcctgg ctgagtattt cccttgcccg gcccacccca    4260
ggcctagact tgtgcctgct gcaggcccttgacccctgac cccattgcac ctgtctccac    4320
aggagccgag gaggtgctgc tgctggcccg gcggacggac ctacggagga tctcgctgga    4380
cacgccggac ttcaccgaca tcgtgctgca ggtggacgac atccggcacg ccattgccat    4440
cgactacgac ccgctagagg gctatgtcta ctggacagat gacgaggtgc gggccatccg    4500
cagggcgtac ctggacgggt ctggggcgca gacgctggtc aacaccgaga tcaacgaccc    4560
cgatggcatc gcggtcgact gggtggcccg aaacctctac tggaccgaca cgggcacgga    4620
ccgcatcgag gtgacgcgcc tcaacggcac ctcccgcaag atcctggtgt cggaggacct    4680
ggacgagccc cgagccatcg cactgcaccc cgtgatgggg taagacgggc gggggctggg    4740
```

```
gcctggagcc agggccaggc caagcacagg cgagagggag attgacctgg acctgtcatt     4800
ctgggacact gtcttgcatc agaacccgga ggagggcttg ttaaaacacc ggcagctggg     4860
ccccaccccc agagcggtga ttcaggagct ccagggcggg gctgaagact tgggtttcta     4920
acaagcaccc cagtggtccg gtgctgctgc tgggtccatg cgtagaaagc cctggagacc     4980
tggagggagc cctttgttcc cctggcttca gtttcctcat ctgtagaatg gaacggtcca     5040
tctgggtgat ttccaggatg acagtagtga cagtaagggc agcctctgtg acactgacca     5100
cagtacaggc caggcctctt tttttctttt tttttttttg agatggagtc tcactctgtc     5160
gcccaggctg gagtgcagtg gtgtgatctc agctcactac aacctctgcc tcctgggctc     5220
aagtgattct cctgcctcag cctcctgagt agctgggatt acaggtgcct gccactgtgc     5280
ttggctaatg tttgtatttt tggtagagat ggggtttcac cgtcttggcc aggctggtcg     5340
caaactcctg acctcaggtg atccacctgc ctcagcctcc caaagtgctg ggattacagg     5400
catgagccac cacgcccggt caggccaggc ctcttttgaa cactttgcac accatgggtc     5460
ttttcatcca ggggggtagg tacagttgta cagttgagga cactgaagcc cagagaggct     5520
cagggacttg cccagggtca cacagcagga tgtggcaggt gtggggctgg gcctggcagc     5580
gtggctccag ctttccagca tagaaatctg tgaaagcaga tagtttgtcg gtcggtaggg     5640
gagactttct gagacccgcc ccagcggctc agagggtagt agccaggggc cttcctgggg     5700
gctcataacc cagaacactg aatgggaaaa ccctgatgga ggaggcgcag tggagctgtg     5760
ggtgccgatg ggaagtccca gaggagctgg gaggtcagta gcggtgctgc cctctgtgga     5820
gcacttagtg ggcaccaggt gtgtttccag gttcatggcc ctgggacctg aagctcagaa     5880
ggtgaagtaa cttgcccagg gcacccgtcg ggcagcggcg ggcagaggat ttgtgggctg     5940
tggagcctgt gctcgtggcc cagccctggg ggttgtgagt gtgctggccg gggagctttt     6000
cctgcaagtg gactggtgtc taggagccag catgtcaggc agcaggcagc gggagtgcag     6060
caggcagcgg gagcacagca ggcagagggc ggggctcgag cagccatccg tggaccctgg     6120
ggcacggagg catgtgggag agggctgctc catggcagtg gctgaagggc tgggttgtgc     6180
cccgaggagg gtggatgagg gtaagaagtg gggtccccag gggctttagc aagaggaggc     6240
ccaggaactg gttgccagct acagtgaagg gaacacggcc ctgaggtcag gagcttggtc     6300
aagtcactgt ctacatgggc ctcggtgtcc tcatctgtga aaaaggaagg gatggggaag     6360
ctgactccaa ggcccctcct agccctggtt tcatgagtct gaggatccca gggacatggg     6420
cttggcagtc tgacctgtga ggtcgtgggg tccagggagg ggcaccgagc tggaagcggg     6480
aggcagaggg gctggccggc tgggtcagac acagctgaag cagaggctgt gacttggggc     6540
ctcagaacct tcacccctga gctgccaccc caggatctgg gttccctcct tggggggccc     6600
cagggaacaa gtcacctgtc ctttgcatag gggagccctt cagctatgtg cagaaggttc     6660
tgctctgccc cttcctccct ctaggtgctc agctcctcca gcccactagt cagatgtgag     6720
gctgccccag accctgggca gggtcatttc tgtccactga cctttgggat gggagatgag     6780
ctcttggccc ctgagagtcc aagggctggt gtggtgaaac ccgcacaggg tggaagtggg     6840
catccctgtc ccaggggagc cccccaggac tctggtcact gggcttgccg ctggcatgct     6900
cagtcctcca gcacttactg acaccagcat ctactgacac caacatttac aaacaccgac     6960
attgaccgac accgacattt accgacactg acatttacca acactgttta ccaacactga     7020
catctactga cactggcatc taccaacact gacatttacc gacactgaca tttaccaaca     7080
ctatttacca acactgacat ctactgacat tggcatctac caacaccaac atttaccgac     7140
accaacattt accaacactg aaatttaccg acaccgacat ttaccgacac cgtttaccaa     7200
caccgacgtt taccgacacc gacatttacc gacactgata tttaccaaca ctgacatcta     7260
ctgacgctgg catctactga caccgatgcc agcatctacc aacaccgaca tttaccaaca     7320
ctgacattta ctgacactga tatctactga cactggcatc tactgacacc aacatttacc     7380
aacaccagca tctaccaaca ccgacattta ccaacaccag catttaccaa caccgatgtt     7440
taccaacgcc gacgtttacc gacgccagca tctaccaaca ctgacattta ccgacaccga     7500
catttaccga cactgacatt tactgacact gacatctact gatactggca tctaccgaca     7560
ctgatattta ccaacgccag catctactga cactgatgtt taccaacacc gacatttacg     7620
agcaccgaca tttactgaca ccaatattta ctgacatcaa catttagcca tgtgatgggg     7680
gccggcttgg gggcaggcct tgctcttggc actggggatg ctgcagagac cagacagact     7740
catggggtca tggacttctg cttcttctcc agcctcatgt actggacaga ctggggagag     7800
aaccctaaaa tcgagtgtgc caacttggat gggcaggagc ggcgtgtgct ggtcaatgcc     7860
tccctcgggt ggcccaacgg cctggccctg gacctgcagg aggggaagct ctactgggga     7920
gacgccaaga cagacaagat cgaggtgagg ctcctgtgga catgtttgat ccaggaggcc     7980
aggcccagcc accccctgca gccagatgta cgtattggcg aggcaccgat gggtgcctgt     8040
gctctgctat ttggccacat ggaatgcttg agaaaatagt tacaatactt tctgacaaaa     8100
acgccttgag agggtagcgc tatacaacgt cctgtggtta cgtaagatgt tatcattcgg     8160
```

ccaggtgcct gtagacacag ctacttggag actgaggtgg gaggatcgct ggagtccaag 8220
agtttgaggc cagcccgggc aaaggggaca caggaatcct ctgcactgct tttgccactt 8280
actgtgagat ttaaattatt tcacaataca aaattaagac aaaaagttaa tcacatatcc 8340
actgccctgc ttaagacaga aaacatgggt gttgttgaag ccagaggcag ctgctggcct 8400
gagtttggtg attggttcct aagcagttga aggcagtttt gtttttccat agatgtctgt 8460
tctccctttg ctgggtgcag cctcgccctg ctgctgtggt cgggtttcag tggcctcgtc 8520
ccgtggacgc agcctcgccc tgccgctgtg gtcgggtttc agtggcctcg tcccgtggac 8580
gcagcctcgc cctgctgctg tggtcgggtt tcagtggcct cgtcccgtgg acgcagcctc 8640
gccctgccgc tgtggtcggg tttcagtggc ctcgtcccgt ggacgcagcc tcgccctgcc 8700
gctgtggtcg ggtttcagtg gcctcgtccc atgggcgtgc tttggcagct tttgctcac 8760
ctgtggagcc tctcttgagc ttttttgttt gttgtttgtt tttgtttgat tttgtttgat 8820
tgtttgtttt tgttgtcgtt gttgttgccc aggctggagt gcagtggcgc gatctcagct 8880
cactgaaacc tctgcctcct tgggttcatg ccattctcct gcctcagcct cccacatagc 8940
tgggattaca agtgcccgcc accacgcctg gctaaatttt gtatttttag tagacagggg 9000
gtttcaccat gttggtcagg ctggtctgga actcctggtc tcacatgatc cacctgcctc 9060
ggcctcccaa agtgttggga ttacaggcgt gagccaccgc gcccagcctc tgttgagcat 9120
attttgaggt tctcttggtg ccagtgatat gtacatgtgt ccccatcgca ccatcgtcac 9180
ccattgaggt gacattggtg cctctcctcg ggtggatgc ctccctctgt ttccagcaac 9240
ttctgaagga ttttcctgag ctgcatcagt ccttgttgac gtcaccatcg ggtcacctt 9300
tgctctcctc agggctccca ggggaggccc gaatcaggca gcttgcaggg cagggcagga 9360
tggagaacac gagtgtgtgt ctgtgttgca ggatttcaga ccctgcttct gagcgggagg 9420
agtctcagca ccttcagggt ggggaaccca gggatggggg aggctgagtg gacgcccttc 9480
ccacgaaaac cctaggagct gcaggtgtgg ccatttcctg ctggagctcc ttgtaaatgt 9540
tttgtttttg gcaaggccca tgtttgcggg ccgctgagga tgatttgcct tcacgcatcc 9600
ccgctacccg tgggagcagg tcagggactc gcgtgtctgt ggcacaccag gcctgtgaca 9660
ggcgttgttc catgtactgt ctcagcagtg gttttcttga gacagggtct cgctcgctca 9720
cccaggcgag agtgcagtgg cgcaatcacg gctcgctgta gcctcaatct ccctgggctc 9780
aggtgatcct cctgcctcac cctctgagta gctgggacta cagacacata ccaccacacc 9840
cagctagttt ttgtgtattt tttgtggggg gagatggggt ttcgctgtgg tgcccaagct 9900
gatctcaaac tcctgaggca caagcgatcc acctgcctcg gcctcccaaa gtgctgggat 9960
gacaggcatc agccgtcaca cgcagctcaa tgattttatt gtggtaaaat aaacatagca 10020
caaaattgat gattttaacc attttaaagt gaacagttca ggctgggcgt ggtggcttat 10080
gcttgtaatc ccagtacttt gagaggctga ggtgggcaga tcacctgagg tcaggagttt 10140
gagaccagcc tggccaacat gatgaaatcc agtctctact aaaaatacaa aaattagccg 10200
ggcatggtgg caggtgcctg taatcccagc tactcgggag gctgaggcag gagaatcgct 10260
tgagcccggg aggtggaggt tgcagtgatc tgagatcatg ccactgcact ccaatctgtg 10320
tgacagagca agactctgtc ttgaaaaata aataaataaa aaaaatttta aaaagtgaac 10380
aattcagggc atttagtatg aggacaatgt ggtgcaggta tctctgctac tatctacttc 10440
tagaacactt tcttctgccc tgaaggaaac cccatgccca ccggcactca cgcccattct 10500
cccctctctc ccagcctctg tcaaccacta atctactttc tgtctctggg ggttcacttc 10560
ttctggacgt tttgtgtgac tggaatcctg caatatgtgg tccctgcgtg tggcttcttt 10620
ccatagcatt gtgttttcca gattcaccca cacattgtcg cacgttatca gaatctcatt 10680
cctgactggg tgcagtgggt taggcctgta atcctaacat tctgggaggc caaggcggga 10740
cgatcacttg aggcaggagt ttgagaccag cctggccagc ctagcaagac cccagctacc 10800
aaaaaatttt aaaagttaac tgaacgtggt ggtggtgggc acttgtggtt cccagctacc 10860
tgggaggctg aggtgggagg atcgcttaag cccaggaggt caaggctgca gtgagctatg 10920
atcgcaccac tgcactccag cctggacaac agagcaagac cctgtctgaa aaaaaaaaca 10980
aaaaaaaaag ttcctttctt tttgtggctg gatgacatcc cattgtatgg ccacagcaca 11040
ttttgtttgt ctgtttatcg ggtggtgggc agtggtttcc accttttgtc tcctgtgaat 11100
aatgctgctg tgaacatttg aattcaagtt tttgtttgaa cacctgttgt gaattatttg 11160
gatatatgtg taggggtagg attgctgagt cctatggtaa tgttaggttt gacttactga 11220
ggaaccatta aactgttttc aacagtggct gcgccgttct gcatccccac cggcagtgtg 11280
tgagggttct gactttacct cctcacaaac gcttcttttc catttaaaaa aatattcagc 11340
caggtgctct ggctcacgcc tgtaatccca gcactttggg aggccgtggc gggcggatca 11400
cctgaggtca ggagttcgag acgagcctgg ccaacatggt gtaaccccat ctctaccaaa 11460
aatataaaaa ttagccgggt gtggcagcgg gcgcctgtaa tcccagctac ttgggaggct 11520
gaggcaggag aatcacttga acccgggagg cagaggttgc agtgagccaa gatcgcgcca 11580

ctacactcca gcctgggtga caagagtgaa actccatcta aaataaaaca aaaataaaaa 11640
taaataaaaa tttattaaaa cattcatcac agccagccta gtgggtgtcc catgtggctt 11700
tgcctcgcat ttccctgata actaggatgc tgagcgtctt gtcccaggct tgccacacct 11760
cagcactttg agatacgtcg cacagtcccc atttgcgaac gagaaatgag gtttagggaa 11820
cagcagctgt gtcatgtcac acagcgagca gggggtctct gagccgtctg accccacagc 11880
cgaccaagct ccaatcctta ccgcctccta gtgttgtgga tgtagcccag ggtgctccca 11940
catttttcag atgagaacac cgaagctcaa aacaggagcg ttttgtccac attggataca 12000
cgatgtctgt ggtttggtcc tgaagtcact ttatatctca gtggtccaga ctggagtagg 12060
acaggggggtt ctggggaatg gggaaggtgt ctcaggtgaa aggaaggaat tccagattct 12120
ccatactgtc cttgggaagt tagaagactc agagggtctg gcaaagtcag acaaagcaag 12180
agaaatgcag tcaggaggaa gcggagctgt ccaggaacag gggggtcgca ggagctcacc 12240
cccaggaact acacttgctg gggccttcgt gtcacaatga cgtgagcact gcgtgttgat 12300
tacccacttt tttttttttt ttgaggtgga gtctcgctct cttgcccagt ctggagtgca 12360
gtggcacgat ctcggctcac tgcaagctct gcctcccggg ttcatgccat tctcctgcct 12420
cagcctcccg cgtagctggg actacaggcg cctgccaccg cgcccggcta atttttgtat 12480
ttttagtaga gatgggattt cactacatta gccaggatgg tctcgatctc ctgacctcat 12540
gatccgcccg tctcggcctc ccaaagtgct gggattacag gcgtgagcca ccgcgcccgg 12600
cccgatttcc cactttaaga atctgtctgt acatcctcaa agccctatac acagtgctgg 12660
gttgctatag ggaatatgag gcttacaggc catggtgctg gacacacaga agggacggag 12720
gtcaggaggt agaagggcgg agagagggaa caggcggagg tcacatcctt ggctttcaaa 12780
atgggccagg gagagacacc ctctgagcat ggtaggacag gaaagcaaga ttggaacaca 12840
ttgagagcaa ccgaggtggc tgggcgtggt ggcttacgcc tgtaatccca acactttgga 12900
aagctgaggt gggtggattg cttgaggcca ggagttcaag accagcctgg ccaacatggt 12960
gagaccccgt ctctactaaa tatacaaaaa ttagccaggc gtgatggtgc atacctgtaa 13020
tcccagctgc ttgggaggct gaggcaggag aattgcttaa acctgggagg cggaggttgc 13080
agtgagccga gatcccgcca ctgcactcca gcctgggcca cagagtgaga ctccatctca 13140
aaaaaaaaaa aaaaaaaaga taaaaagacc aaccgaggaa ttgaagtggg ggggcgtcac 13200
agtagcagaa gggggatcgt ggagcaggcc accctgtggt catgcactgg aagctcatta 13260
cctgacgatt tggagctcat cactggggggc ctaaggagaa tagatactga aggatgagga 13320
gtgatggcgc ggggcacggg tgtctttggt ggccagaact tggggactgc tggggtgcct 13380
cactgcaggc cttctcagcg ccctttatat gcttacacag gctgtttcta agagggggat 13440
acattgcata agcgttttca gactacctca tcatgggtcc ctttctttac cctctgtggc 13500
cctggtggcg cactctctgg gaaggtgcag gtggatgccc agacccgccc tgccatccac 13560
ctgcacgtcc agagctgact tagcctcgag attgctgctg gcacctcctg ccccgggaca 13620
cctcggatgt gcccgtggag atgctggctc tgtgttttct gctggagttt ggtgcgtctt 13680
ttcctcctgc aagtggccac cgctcttggg tatgtcctca ggcttctgcg agtcatggct 13740
gcttctcagg tccttgccca gcgccaggag caaaccctcc tggcactttg ttcaggggtg 13800
gatgcgccag tgttcctgct gtggaccccc atctcacatg agggtcttgg gcctgcaggc 13860
tcgttcagga aacacccgct gagtacgcag tgtgtgccag ctgtgtccca ggcaatggcg 13920
gggacagtgg ctgctgctgg ggttgtggtg gcttctgggg actctgggga cagctgaggt 13980
gcaaggagcc acggctcctt gaggatgcag ttggactcca ggtggaaggg atggttgggg 14040
gaggtataaa tggggtcagg gaggagacac atttggaaca atgggaacat ttttaagatg 14100
ctatgtcggg aggcaacaag gtggccaacc caggtgctga ggagcccaca ccagccctgg 14160
acgtgttttg ccgctcacct ttgctgggga gtggtgggag agaggattcc gttccacgtg 14220
gtggtgtgcg cagctgggct gtgtggagct gggcgctagg aggaaggtgc tttctgcggg 14280
gctagccggg ctctgccttt gaacacaatc aggctccagg ttttcagcat ccagtgcatg 14340
agaggacttc acgggcagct gtggctgatc ccttgatgaa ttgggagaag aacaaaggtc 14400
tatgaaatga ggtttcatgt agatggcatt agagacgccc acaacagatt tacagagtgg 14460
agcggagacg gcggatgggt ctgggaggcc cctcctgctg gccttgactg tgacagctgt 14520
cctgggaatc agcttccagg ccgcccccagc agcctgactg acacacacag gggttttagc 14580
cccatcctgc gaccagctgt tgccatcatc agtgacagct gggagtggcg gtggttccag 14640
ccctgggcac cctccccacc tgctggggcc cacccagggc agtcctgaca cctacaggtt 14700
gcttggagcc gcatccgagt cctgccccac cacgtgtgaa gcccgagtgg tcgtgggctg 14760
aggtcccctg attgcatccc cacttccctt ctgcttcaca tagctgcctc ttctcaccgt 14820
ttttccagcc tcctgggcta ggaattccag tgttgtgctg gctttgcccc aggacacctc 14880
cttagccctc ttcctgagtc tagagccccg gggttggaa gttctggccc ctgggacacc 14940
tgcagccaca ctcagcttct cctgtgagcc tccagcatgt cccctcagga ccaagccctc 15000

acgttcttgc ctccccgccc acctgggctc agccagggga aggcctggct gggagcgtct 15060
cccctctgcc ctgcccttct cccctctacc ctgcccttct ctcctctgcc ccgccatggc 15120
ttttatatcc tgtgccacaa gacatggctg tgtgtgaaag tggcagggtc tggcatctct 15180
gtgggtctct gaggcccacg ctccagtgcc actcttccca cccgctggcc gtgccctcat 15240
gctggaggga cagcccagcc ctctcccgaa ccccagcccc atgtgcccag ctgcccccgg 15300
ccctctcccc tggaagccgg ggtcactcca gccgtatgcc atggtgggga catcctgctt 15360
ccttggcctt ccagggaagg tcctctttcc aaatggcgac acctggtccc tgcctggagg 15420
ctggaagctg tggccccttgt atgcccctcc agggtctgtg cgctcggttg gcccgagttc 15480
ccatcaccgt catcatcacc atcatcattg tcatttcgct tgtctgtgag ccggcctggt 15540
ctcccagagc agagaccctc tgaggtccag cctgagttgg ggtctccgtg ctgacccctg 15600
acggggactc aggacgtacc aggtctgggt caggagtgac ccccaaacct cgtgcccttt 15660
gacaggcacc cctgactttt gctaagtggg tggaggtgac atcacttaca gcgggagtga 15720
tgggacaggg tctgttggct gcactgtgct cccagggatc tggggagagg ctatatccct 15780
gggctttggc actgcagagc tgtgtgtgtt tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 15840
tgtgtgtgtg tgtgtgtgtg tttgcgtgcg cgcacatgtg tataagatct ttttttatta 15900
catgaagcaa gataactgtt gctgtttcct tttgggtttt gtgttcaaca gagtggggta 15960
cttcttccct cagacaacag aactctcccc tttaaacacg tgctgtcaga gggtgggtct 16020
tgggctcatg tctgtttgca cagccgagtc agaggaaaca cagggttctt cataaaaaca 16080
ctgcacagca ggcgactgtc cagagtcagc ctgcaggacg gcagcagccc tgcccctcag 16140
agcacagcta gggtgggctg ctttgggatc tcccgtcatt ccctcccagc tggcagccgg 16200
cggccggccc attccttggt gtgctggtca ggggggcgtg cgcctgctct gctcaccctg 16260
ggaatgggac agaagctggc agctcggaga ggacagggct ggacccttgg gtggcctctg 16320
gctggaccat ctcattgtcc tcagacacag cctctcgggt ctagtttcat ttcctgaaaa 16380
acaagtgcac agaactagag caggagtcga gagctacggc ccccgggcca gatccagccc 16440
tgccacctgt tttcacacca tgctcaagct gagtgggttt tacatttttt aattacttga 16500
aaaaaaaaaa gccaaaggag gtttcatgac ccatgaaaat tatatggaat tcaaaaaaaa 16560
aaaattatat ggaattcaaa tttcagtgtc cataaaataat ttcttgagac agggtctcgc 16620
tctgtcaccc aggctggagt gcagtgctat ggcatggctc gctgtaccct tgacctccca 16680
ggctcaagcg atcctcctgt ctcagcctcc tgagtagctg ggactacggg tgtgtgccac 16740
caagcccggc taattttttt ttaattttag taaagacagg gtctttctat gttgcccagg 16800
cttttctgga actccatctt ggcctcccaa agtgctggga ttacaggctc gagccacgga 16860
gcccagcctg tttttgtttt ttcactgata aagttttgcc gggtgtggta gtgtgtgcct 16920
ctagcgattt gggaggctga ggtgggagga tcgcttaagc ccaggagttt gaggctgggc 16980
tcaagtgatc aggaggtgaa ctatgatcat gtcattgcat tccagcctgg gtgacagagc 17040
aagaacctat ctcttaaaaa tatatattta aaaagtattg ggtgtggtgg ctcacgcctg 17100
tggtcccagc tacttaggca tctgaggtgg gaggatggct tgagcccagg agtttgaggt 17160
tgcagcgagc caagatcgtg tcactacact ctagcctggg tgacagagcc cagaccctgc 17220
ctctttaaaa aaaaaaacca aaaaacatgt attggaacac agccatgcct gttcagtcac 17280
gtgctctcca tgctgctttc tgctccagag acccttatgg cctgaaagct gaaaatattt 17340
tctatccttt acaaaaaagt ttgctgacct ctgtcctgga aaattcatct cccaagttct 17400
cttccggcac tggcgttcct gggtgtccta aatttggccc ctgttatttc tgaactctgt 17460
tttggctctg ttccctccca ggagccagga caggcacgtt ctctgcatct tgtcccctga 17520
cgcccagagg cttggctcgg ctcaggcatt cttggaaata tctggctcca ggaaaggcag 17580
aggcctcctg agtcagccca gagggaacct gccccaggtc tgggggaggc ctgacccagc 17640
agagtggctt ttgccgatgg gttgggccgg tcaagatgtg ctgaaagttg tcctcagaag 17700
gccactttgg gattccttcc tccagtatta gagcaactga gagctgctca ttgcaagcct 17760
gatgttttcc cagttggccg ggtccaccgg gtgccctggg attctgggat ctgggtggaa 17820
agtaggggc ttgggggagt gtcctgggtt ctggaatcca ggtggcaagt ggtgaggttc 17880
agggagtggc ttctgagcca cccataggggt ctctgtggga ggctctgccc atccaggaga 17940
ttccgcaggc cctgccggcc cagagccagc gtcttgcgct tgccgaggct acagccagcc 18000
ccagccgggt ggaacagccc gtcgcctcct ctcactttgt tttggggcca cctgggagtg 18060
tggagcaagg gtagagaggg aggaagtggc tgccggccgc tgcccagcac ccttgtttgc 18120
cttgggccct ctgtgggctc ctttttattg ctcttcaatg aagccaggga aatggacttc 18180
cttgcctcac ttcagttcaa catgtctgga agtttggtat taaaattaag aaagtgtgga 18240
aatagagcaa gaagagaaaa atctctccaa gagataatag tgacctctga gctgggcgcg 18300
gtggctcacg cctgtaaatc ccagtacttt gggaggctga ggcgggcaga tcacctgagg 18360
tcgggagttt gtgaccggcc tgaccaagat ggagaaaccc cgtctctact aaaaataaat 18420

```
aaataaataa ataaataaat acaaaattag ccaggcatgg tggcgcctgc ctataatccc    18480
agctaaggca ggagaatcgc ttgaacctgg gaggcaaagg ttgcagtgag ccaagatcac    18540
gccattgcac tctagtctgg gcaacaagag tgaaactccg tctcaaaaaa aataaataaa    18600
taaaaaataa aaatagtgac ctctggccag gtgtggcagc tcatacccgt aatcccagca    18660
ctttggaagg aaggccgaga tgggcagatt gctttagcac aggagtttga gaccagcctg    18720
gccaacatgg tggaacccca tctctacaaa aatagaataa aatttaagag gtaatagtga    18780
ccttttggta gatcgaaacc tggattgctt tctttttcta aatgctgatt cttttctttg    18840
tggtgtttgt gttctgtgcc gatgtccctc ccccagccct gttattgtga gtggaagaag    18900
gggaaagggt tcgcccgcta ctgtgagccc ctcctctcac gctgggtgtc cttggagaag    18960
cctgcacttc ttcattgtac gccagggctg ggtccctccc tggagtggtt ctgtgctgct    19020
gggatggggc caacccctca gatgttttct gagtgtcaca cacaggtgtg tgcattcatg    19080
gcctttgcgt gtcttcctgt tgtggaggca aaaatgtgaa gaaccctaga tgattttggg    19140
accagggctc catcacctgc tgttcattgc acaccggagc atccaggcat gggtggagag    19200
ctcagacttc caggcacggt cgcaggggct ggtctaacca tgttcccgcc cgcctgctcg    19260
tcagaaccgc ctgttgggag ctgttatcat gataccatac ctgggccctg ggctatccga    19320
ttctgactta attgctccag gttggggcca ggccgttgtt tgctgttttg ttgtttcttc    19380
tgtgacgtta gccactgggc taatctgagc ccctcagtta caggtggaga aactgagacc    19440
catgggggtg caaggacttg ccgaggaccc agagcccctt gggggcagag ctgaggcggg    19500
gcctggcttt gggtcccaga gcttccagtc cccttcccgc tctcctaaca gctttttttt    19560
ttgagacaag atctcaccct gtcacccagg ctggagtgca atggcatgat ctcggctcac    19620
tgcaatcttc gctagctgcg ttccagcgat tctcctgcct cagcctcccg agcagctggg    19680
attacaggtg tgtgccgcca tgcccagctc gttttttttt gtacttttag tagagatagg    19740
gtttcaccat gttggccagg ctgatctcga actcctgacc tcaaatgatc cgcctgcctc    19800
ggcctcccaa agtgctagga ttacaggctg ggatcacact gtgcctggcc ctagcagctt    19860
tgtcctgtgc catccaacaa cagatgaccg aagtctttgt ttcttaacat gcattccatc    19920
tgccttacag ttttgccacc tgcaaaacag aggacttgtc gcttttctgg taagctggaa    19980
atgtaatctg gtagcaggag gcctgtggaa gcttgccttt aatggccttg tgtctctttc    20040
atcctgtcct gagagccgga gaacttggat gttgcaccta actcaacctt cctgttaaca    20100
tacagttctg caggctcatg gatcatcaga accacgtcct atctcacgcg gctgtatgct    20160
tccgttggtt caggtgtttt taccttgaca gtatttctc ctcggtggct tttgcggtgg     20220
ttgcttttaa tcagcattga ctcttcaaga aaaatattta gctgctacat ctcagaggag    20280
acagggtgga aagcatctga gacctgcagg ctcagactta gaaccagaag tgccctcaga    20340
gttcatccgg ccctgaccca gcgggaaatg agttcacaga gaagcgggag aactttgccc    20400
caggccctgc cgttgctcat aactgcccca ggtccttaca tttgctccag gtcctgcccc    20460
aggccctgca gttgctcata actgccccag gtccttatat ttgctccagg tcctgcccca    20520
ggtcctgcag ttgctctgtg tggtgggtgt gatctggagc cctccgccca ttgctgcacc    20580
tggggcaggc attgctaatt gatcccagga ctccttcctg cggagcacgc cctggttctc    20640
caggcagccg ctgcctgtca gcctgcagtg gttcgggaga ggacacctgc ttgcctggtc    20700
tgttccaaat cttgcttctc atcccagcac aggtaggggg tgctatggga aagggatcct    20760
cagttggccc tgtcactgct ctatcagctg ggacgtggc atcctagtga aaacatcatg     20820
gccgggcgcg gtggctcacg cctggaatcc cagcactttg ggaggctgag gagggtggat    20880
cacttgaggt cagaagttcg agaccagcct ggtcaacatg gtgaaaccca tctctactaa    20940
aaatacaaaa attcgccagg tgtggtggcg ggtacctgta atccgagcta ctcgggaggc    21000
tgaggcagga gaatcgcttg aacctgggag gtggagcttg cagtgagccg agatcttgcc    21060
actgcactcc agcctgggca acagagtgag acgctgtctc aaaatctcaa acaaacaaac    21120
aaacaaaaaa caaacaaaca aagcgtcatt tatccagcac ccctggggaa ccatgctacc    21180
tggtgtttta tggtacctgg caaggtgcag gtgaagttgc tgctcttggg cattgaaccc    21240
gtcttgtttg gggcagctca ggccccaggc agggtccggg ttggctctcg ttggtgtggc    21300
cctggcccat ccagacctat atttctgccg tcctgcaggt gatcaatgtt gatgggacga    21360
agaggcggac cctcctggag gacaagctcc cgcacatttt cgggttcacg ctgctggggg    21420
acttcatcta ctggactgac tggcagcgcc gcagcatcga gcgggtgcac aaggtcaagg    21480
ccagccggga cgtcatcatt gaccagctgc ccgacctgat ggggctcaaa gctgtgaatg    21540
tggccaaggt cgtcggtgag tccgggggt cccaagccat ggctcagcca tgcagacttg     21600
catgaggagg aagtgacggg tccatgcctg ggcataagtg ttgagctcag gtgccccgac    21660
ctggggaagg gcaggacagg aaaggtgaca gtatctggcc aaggacagat gggaagggac    21720
caagggagct gattagggag tggttatgga ctaggaatgt cggtaacaat ggttagaaag    21780
tgactaacat ttgttgagca cctgctgtgt gcccggccct ggccgggagc cttcgtgccc    21840
```

acagtgaccc cgtctgcaaa tgtagttcct tgccctactc gcactgggga gcaggacgca 21900
gagccgtgca tctcacaggt gccaagctca ggactccctc ctgggtctgc ctgggctggg 21960
ctgtgcttgt tgcccctgtg gcccacgcat gtgcaccttc cacctgaaag ccaggatctt 22020
caggacgctc cccgaggagg tcgttgtctg gcacaatgat ttgtctcttc ctgaaaaggt 22080
gacagagtta cactggagag agcagcatcc aggtgcggca gggacaggcc tggggctcgc 22140
gggcagggac tctgtgtcct gccggggtcc cacactgcac ctgcttgtca gaggcactca 22200
gtcaatcttt gctgatgaag gatgagagga cagaggacgt gatgcttgct gctgcattgc 22260
ctgcagtcct gggtgagatg cccgggttga ctctgctgcc cgtcgggtgg atgtgatgtc 22320
agatccccgg ctttaaaata cgagggagct gggaattgag ggagcaggtt ggggcagaaa 22380
gcacagcccc gtggaagcct ggagctgagg cagtgtgggc gacccctgga gcagtgagtg 22440
cttccttcat ggccttcatc gcaccctgca gtcctcatgt aggggatgcc atccatgaat 22500
ttagttttcc cagcctcctt taaaaacgcg ttcatgctgg ggccggggca gtgcagtggc 22560
tcacatctga aatcccacca ctttgggagg ccgaggcggg tggatcatga ggtcaggaga 22620
tcgagaccat cctggctaac aaggtgaaac cccgtctcta ctaaaaatac aaaaaattag 22680
ccgggtgcgg tggcgggcgc ctgtagtccc agctactcgg gaggctgagg caggagaatg 22740
gcgtgaaccc gggaagcgga gcttgcagtg agccgagatt gcgccactgc agtccgcagt 22800
ccggcctggg cgacagagcg agactccgtc tcaaaaaaaa aaaaaaaagt acaaaaaaaa 22860
aaaaattagt ctgggtgtgg tatcacgcgc ctataatctc actactcgag aggctgaggc 22920
ggagaattgc ttgaacccag gaggtagagg ttgtagtgag cccgtatcgt accactgccc 22980
tccacctggg caatagagcg agactctgtc tcaaaaagaa aaaaaaaaaa agaacattta 23040
tgccaggtgt ggtggctcat gcctgaaatc ccagaacttt ggaagactga ggcaggagga 23100
tcacttgagc ccagaaattt gagagtgtct tccctgggca acatagagag acctcatctc 23160
taccagaaaa aaaaaaatta gcccggcatg gtggcatatc cctgtggtcc cagctactta 23220
gggggctgac gtggcaggat cacctgagtc tggaggcaga ggttgaagtg agctgagatc 23280
atgccactgc actccagcct gggtgacaga cagagaccct gtctcaaaaa aaaaaaaaaa 23340
aaaaagcatt tactatccac catggaaggt gagactgacc tgtgagtgat tgttcaaaga 23400
acaaaaaata aaccccagag ataagacaaa agggtgcctc catgggggtg tgatttaaag 23460
ctgagaaatt gggcttcttc cccctcccct ctcaccccgt ggtttgctaa aggagatggg 23520
aaaaaggatt cttttttttgg ctgaaatatt taacactaaa ttaaagccaa ttttaacagc 23580
actttggttg atgagtgaaa ttaacagact ggccaaaaat aaacgaacgg tctgtactat 23640
gtgaaaaaga ggcagctttg gccatgctgg gccaatgtga gttttcaggg ttgctgggaa 23700
tgtctgtgaa tcggaggaag ggcctagctg ggactctcag gagccaaggc cctgaggggc 23760
aacttgcctg gtccctgccc tgaggcgttc actgctttct tcctgggcca gatcacaggc 23820
ccggaggctg gaccactggg ctggcactct tgccgagctg ctccctgact tcctgaccat 23880
gctcctttca gcagccttgc tgcactttag tttccttgaa tgaaaaatgg ggatgagaat 23940
agctcctacc tccaaggtga atggagtgag ttcggacagg tgactccctg ggaccagtgc 24000
ctggcgcctg acaaggtcca gtcagagccc gcactgctgt tactgatacc cttggctgta 24060
ccaggggaga acttggttgc cattgccagg tgttctccca ccaccccccac tactgtccct 24120
gtttgatgtg tggcgggaat aaagctgtgc acattggagc ttttggcaca tcctggcttt 24180
caggtgaaag gtgcgtgtgt gtttgagggt ttagcctggc caacccagcc atgaggtcgg 24240
acctgacctg gggtgagtc ctgagctcgg cacccctgag ctgtgtggct cacggcagca 24300
ttcattgtgt ggcttggccg caccccttttc cctgctgggc tgttgatgtt tagactggag 24360
cctctgtgtt cgcttccagg aaccaacccg tgtgcggaca ggaacggggg gtgcagccac 24420
ctgtgcttct tcacacccca cgcaacccgg tgtggctgcc ccatcggcct ggagctgctg 24480
agtgacatga agacctgcat cgtgcctgag gccttcttgg tcttcaccag cagagccgcc 24540
atccacagga tctccctcga gaccaataac aacgacgtgg ccatcccgct cacgggcgtc 24600
aaggaggcct cagccctgga ctttgatgtg tccaacaacc acatctactg gacagacgtc 24660
agcctgaagg tagcgtgggc cagaacgtgc acacaggcag cctttatggg aaaaccttgc 24720
ctctgttcct gcctcaaagg cttcagacac ttttcttaaa gcactatcgt atttattgta 24780
acgcagttca agctaatcaa atatgagcaa gcctatttaa aaaaaaaaaa gatgattata 24840
atgagcaagt ccggtagaca cacataaggg cttttgtgaa atgcttgtgt gaatgtgaaa 24900
tatttgttgt ccgttgagct tgacttcaga caccccaccc actcccttgt cggtgcccgt 24960
ttgctcagca gactctttct tcatttatag tgcaaatgta aacatccagg acaaatacag 25020
gaagactttt ttttttttt tttgagacag agtcttactc tgttgcccag gctggagtac 25080
cgtagcgtga gctcagctca ctgcaacctc cgcctcccag gttcaagcga ttcttctgcc 25140
tcagcctcct gagtagctgg gactacagac atgcaccacc acacccagct aattttttt 25200
atatttttag tagagacagg gtttcatcat gttggccagg ctggtcttga actcctgacc 25260

```
tcaggtgatc tgcccgcctc ggcctcccaa agtgctgaga taacaggtgt gagccaccgt    25320
tcccggcata ggaaaacttt ttgccttcta aagaagagtt tagcaaacta gtctgtgggc    25380
tggccttctg attctgtaaa gaaagtttga ttggtggctg ggtgcggtgg ctcacacctg    25440
taatcccatc actttgggag gccgacgtgg gcatatcacc tgatgtcggg acttcgagac    25500
cagcctcacc aacgtggaga aaccccgtct ctactaaaaa tacaaaaaaa aaattaaccg    25560
ggcatggcgg cgcctgcctg taatcgcagc tactcaggag gctgaagcag gagaattgct    25620
tgaacctggg aggcggaggt tgtggtgagc tgagatggca ccattgcact ccagcctggg    25680
caacaaaagt gaaactccgt ctcagaaaaa aaaaagtttg attggtgtaa ccaaagcgca    25740
tttgtttatg gattgtctgt ggcagctttt gttctgccga gatgagttgt gacagatctg    25800
tatgggctct aaagcctaaa acatgtgcca tccgcccctt tacagaaaaa gtgtgctgac    25860
ctctgttcta aagtattgga caactacaat gtttgctcat ttattattct atgatttgtt    25920
ttctgctttt tgttgttgtt gttgttgttg agatagggtt tccctctgtc actcaggctg    25980
gagtgcagtg gtgtaatctc agctcactgc agcctcgacc tcctgggctc tagtgatcct    26040
ctcatctcag cctccctagt agctgggact acaggcacac accaccactc ctggctgatt    26100
tttttttttt tttttttttt ttgtggagac agggtttccg catgttgccc aggctggttt    26160
caaactccta ggctcaaaca cccacctcag cctcccaaag tgctgggatt acaggcgtga    26220
gccaccatgc ccagcctatt ctactgtttg tattacatag ctttaaaaga tttttttatga    26280
ctttaagtca caagggttct ttgtagaaaa aaatatatat ataggaaagt ataaaaagaa    26340
agtaaaaatt gtccataacc tctccagcca gagacgaccg ttgctgacac ctcagcatat    26400
tgcctttaag tcttttttct ctaagatagc atttctcttc atcacagtca tatgctacgc    26460
agaattctgt atcctgattt tttcacttga cattacaaca ggtatttgat ggcgctgtga    26520
caaactcttt ggcacaatct tttaaatgta tgaaatactc cactgcacag atgtttgctt    26580
ttaggcttaa ctgttctttt attttgcgtg tgctggttac agccgggcac agtggctcat    26640
gcctgtaatc acaacacttt gagagggtga ggcaggagga tcacttgagc ccagaagttt    26700
gagaccggcc tgggcaacat agtgagaccc catctctaca aaaaactttt ttaataagtc    26760
gggcgtagtg gtgcatagct gtagtcccag ccaccaagga ggctgagttg ggaggattgc    26820
ttgagcccca ggaggttgat gctgcagtga cctgagatta ctccactgta ctccaacctg    26880
agcgacagag caagacttgt ctggggaaaa aaaaaaaaaa aatatatata tatatatata    26940
tatatataca tatatacata cacgcacaca cacataatat aaaaatatat atttataaat    27000
atataatata taatataaaa atatatattt ataaataaaa tttataaatt atatttataa    27060
gtaaatatat aatatataat ataaaaatat atattatata atatataata aaatatataa    27120
tataaaaata tatatttata aataatatat aatacatact tataagtata tatttaaaat    27180
atatgtaatg tatatttttt aatgtatgat atataatata catttataaa tacacattta    27240
tattatttta tataaaatat atataaaatc tccaagttgc tttttccaaa aaggtgtctt    27300
gctgcatttc aaacattcat ttaaaaactt gaatgctggt gatctggtcc agaatgtgtt    27360
cagtagctgc tgccagtggc caagcatctc gggagatgtc tacaaaacac gctggttctg    27420
gcctggcgtg gtggctcacg cctgtaatct cagcactttg ggaggctgag gcaggtggat    27480
caactgaggt ctggatttcg agaccagcct tgccagcttg gtgaaacccc atctctacta    27540
agaatacaaa aaaattagcc aggcgtggtg gcatgtgcct gtaatcccac ctacttggga    27600
ggctaaggct ggagaatcgc ttgaacccag ggggcagagg ttgcagtgag ccgagatcgc    27660
accattgcac tccaggctgg gcaagaagag cgaaactccg tctcaaaaaa aaaaaaaaag    27720
atgctggttc ctaaaatgtg gccctttccc tcctcacctg ctgccagacc atcagccgcg    27780
ccttcatgaa cgggagctcg gtggagcacg tggtggagtt tggccttgac taccccgagg    27840
gcatggccgt tgactggatg ggcaagaacc tctactgggc cgacactggg accaacagaa    27900
tcgaagtggc gcggctggac gggcagttcc ggcaagtcct cgtgtggagg gacttggaca    27960
acccgaggtc gctggccctg gatcccacca aggggtaagt gtttgcctgt cccgtgcgtc    28020
cttgtgttca cctcgtatga gacagtgcgg gggtgccaac tgggcaaggt ggcaggctgt    28080
ccgtgtggcc ctcagtgatt agagctgtac tgatgtcatt agccttgatg gtggccagga    28140
ctggtagggc cctcagaggt catggagttc cttcgtggag cgggtgctga ggctgtatca    28200
ggcacagtgc tggctgcttt cacctgggcc gtctcaccga agtgtccatg gagcctgcgt    28260
agggtgggta tctgtgtcga ttttacagat gcagaaacag gctcagagaa accgagtgac    28320
ttccctaagg tcacataccc agttagagca gagctgggcc aggaagtgct gtctcaggct    28380
cctgaccagg tctccttgct ttgcactctt gccaaaacca tgatccagaa ctgactttga    28440
ggtccccgga cctcaggctc ctccgaaatg gcctcttgga ggctgctgag ccacagctta    28500
ggacccacct cgagaggcaa atgtgctttg agctgccagg cgtcctgggg gccctgcctt    28560
gggcacgggg ttcagacagg ccccagatgt gtggggcgtc tttctggact tgagttttct    28620
tttctgtgtg gtggacacag tgctcacccc ttaaagcacc tgtgatgtgt gcagcagccc    28680
```

aatccctgcc tgtcgcctgt tctgctaggg aaggaaggaa gacttcagga tggcaggaca 28740
acagaaagag gtccaggttt tagagcaagg gcaggtcaaa cttagaaaat tctggaatga 28800
ggatgtgcat ttcctcttct ggatctgcta aaagaagagg gaaggagggg ctgctggggg 28860
aggagcccag agccgagttt acatccggat cccgcaaggc ctcccctgcc ctgaggtctt 28920
gttttgtgat gtgcttgtgt ccatcctggt ttctgccgtg tccccaacat ccggccaagc 28980
ttaggtggat gttccagcac acactcaccc tgtctgtgca cctgtttttg tgtccgtaag 29040
tgggtattta ctcaccttac gagtgagcca ctgtgggaat tcagggaggt ggcgcagtga 29100
ccacccctgg agggatatgt gtgtggcagg ggtcgagggt ctcgcccttc cctgcttcct 29160
gcgcgtggct ttctccagga cggggagggc tgagctgaag aggtggggac agttgcgtcc 29220
ccccgccacc cactgtcctg cggtgagagc agactcactg agcctgccct tctcccttgt 29280
gccttccagc tacatctact ggaccgagtg gggcggcaag ccgaggatcg tgcgggcctt 29340
catggacggg accaactgca tgacgctggt ggacaaggtg ggccgggcca acgacctcac 29400
cattgactac gctgaccagc gcctctactg gaccgacctg gacaccaaca tgatcgagtc 29460
gtccaacatg ctgggtgagg gccgggctgg ggccttctgg tcatggaggg cggggcagcc 29520
gggcgttggc cacctcccag cctcgccgca cgtaccctgt ggcctgcaag ttccccaacc 29580
tggcaggagc tgtggccaca cccacgactg cccagcagcc tcaccctctg ctgtgggagt 29640
tgtccccgtc cacccctggg tgcctttgct gcagttatgt cgggagaggc tctggtgaca 29700
gctgtttcct gtgcacctgc tgggcactag gtcccagcta atccctgtgc caggactcta 29760
atttcaccct aacacacatg gtggttttca ttgctgggga agctgaggcc tgagcacatg 29820
acttgcctta ggtcacatag ctggtgagtt caggatcccc cagagatacc agggccagca 29880
ctcgatcccc acccagccct gaaccccacc atgtgctggg attgtgctgg gagtgtccac 29940
acgcctggga ccccagggct ggtgctctca tctccttttt ccagatcatg agaatgaggc 30000
tcagggaagt ttgaaaaaaa cctatcccaa gtcacacagc aacaggagca ggatttgaac 30060
ccagaaaagg ggaccgcaca ctctgttctg ctagagtagt tagctgtcct gggtgatatg 30120
gcaggtgaca ggggcaactg tgcttaacaa aggaaccccc atcccccctg ccaagttggg 30180
agactagaag gtcaggggca gaagctctga agggccaggt gcagtggctg acacctctaa 30240
tcccagcact ttgtgaggcc aaggcgggca gatgatttga gcccaggagt tcaagatcag 30300
cctgggtaat gtagtgagac gccatctcta caaaaaaatt ttttaaaaat tagctgggca 30360
tggtggttca tgcctgtagt ccaagctact tgggaggctc aggtgggagg attgcttgag 30420
cccaggaggt tgaggttgtg gtgagctgtg atcatgccac tgcactccag cctgggcaat 30480
agagtgagac cgtctccaaa aaaaaaaaaa gaagaagaaa aagaagctct gaggctccaa 30540
gtccccaggc acccccttggc ttgagggcag acaagggagg agagggtcac ctgggcagcc 30600
ctgacttttg tcccctggca aagggacctt cagtgacctt ggccctagga gagcctctga 30660
gcacgtcagc catgtcgaac cgctcaggaa gggcagcaag aatttggctt ctgacctctg 30720
cctctcctac tcgccatctg cactgggtgt ggttgtgccc attttacaga tgaggaggct 30780
ggggcatcga ccagctgaat gccttgtccc aggtactgcg taggcagagc tggcagttga 30840
accccgtgtc ctggttgtcg ctgggggtgg gctgcaccct gacttgtgag gccagtagca 30900
aggtttgcac gtgacttcgt gaccgtcacc cagctctgca gcacatcccg tgacccagct 30960
catccaggcc gcatgcaaac ctgttgccag gcgagaaacc agtcaccgca cagctgtggt 31020
tgcctgaaat gattaagctc attaatcacc ccggagtgag gacagactca gatgaaaacc 31080
agcaaaagcc ctggaaactc atgtgaccct gccaatgagg gcggccatgt gcattgcagc 31140
ctggccgtca ctcctcggta cgtgttttgg acttaaacgc tccggatgtt tactgagtgc 31200
ttgattaata acatggaagg cctggtctca ttgctgtggg agtgaaggat gcacagccag 31260
gcctgacatg atgagaacaa gaacctggag tctcgctgcc tgggtggtaa tcctggccct 31320
gccacttagc aactgtgtga ctgtagccag gtcacttaat tttgctagat cctgcctgcg 31380
cttcagtgga tcttgctggt tttccaaggt ggccaaacac tttaaggcat tcatgtggtc 31440
gctaggctgc agggttgaac cctggctcac cccgcagggc gccgtgtgct ctgtggcctg 31500
gctgtgcctt tgctgacacc gtgcccgtgt gtgttcatgc aggtcaggag cgggtcgtga 31560
ttgccgacga tctcccgcac ccgttcggtc tgacgcagta cagcgattat atctactgga 31620
cagactggaa tctgcacagc attgagcggg ccgacaagac tagcggccgg aaccgcaccc 31680
tcatccaggg ccacctggac ttcgtgatgg acatcctggt gttccactcc tcccgccagg 31740
atggcctcaa tgactgtatg cacaacaacg ggcagtgtgg gcagctgtgc cttgccatcc 31800
ccggcggcca ccgctgcggc tgcgcctcac actacaccct ggaccccagc agccgcaact 31860
gcagccgtaa gtgcctcatg gtcccccgca cctcactccc tcgttagatc aggctggttc 31920
tgggagctga cgctgaaagg agcttctcat ctggggttcc tgggtgtaca tagatggttg 31980
ggtaggttgt gcactgcaca agctgcatga tgctacctgg gggtccaggt ccaggctgga 32040
tggacttgtt gcttcatcag gacatagata aatggccaaa actcctcagc tggaaggtcc 32100

tgggcaggat ctttgggtgt gaaaaccagt cacaggggaa gggtgcttgc tcatactgcc 32160
agcacagtgc tgagtgcttt ccatagcgct cgtttactcc tcaagcctgg agggtgggga 32220
gtagcatggt cccatttcac gtacaaggaa cccgatgcac agagaggtgt ggcaacccat 32280
ccaaggccat acaactgggg tgggttgagc cggggttgac tgtggcaggc tggctcaaga 32340
gtccctgctc ctgaaccctt gccaggcagc ctggcatcag ctcggggaat ttttgccctg 32400
acccttggaa gcaagtgggc ctctttgttc tcatgtcagt gatgagaaga gtgactttcc 32460
tatggcccct ctggagtaca ggtgtttcct gttggcgggc tcttccccca tgacatcagc 32520
agcgagctgg ttatgattcc ctacgcagaa cttgatagtt tataaagctc tttgtcatcc 32580
aggccccgtt ggagtctcac gcagacctgg tcgcaggcgg ggctggtctt gcctgtccca 32640
gctgcatgga tggggaactt gaggcttgca aaggttaagg ggctgttcga ggcccaggct 32700
ggcaggagat gggcctgggc cagagtctgg gacttcccat gcctgggctg tctttggtcc 32760
tgttgctcac catccctccc tggggccatg accttagaga gccaaatgga ggtgcaggta 32820
acccacggca aggaggggtt gccatgactc agagtccccg tcctgtggcc ggcagtacct 32880
ggtgcaacga cttggatttc agaccagcca ctgtagcccg ctgacggtgc gctcgaagtg 32940
ccacagcttc tgaagccagg caggactcag gccaggagac tctgttagct gttgagaggg 33000
agaggccaac ggatgttctg gttctgctag agagctggtt cttcggatcc tggtaccagt 33060
gcactgagag gaggcccagc ttgattctgg ggctgccttg tggtggcatg tgctgctcac 33120
tgacaccctc gaggagtgtc ttctctcggg cttgttgact gtgcccggtt ttccgcagtt 33180
cactggtgca cacataggca catagcaaac cgcacacaca gtcgtgggta tgagtttcac 33240
tacattccac caccagtgtt cactaccatt acctgccttc cgtcttaagt gttcatcatt 33300
taaaaataaa tttattgggc tggacgcggt ggctcatgac tgttatccca gcactttggg 33360
aggctgaggc gggcagatca cctgaggtca ggagttcaag accagcctgg ccaatatggt 33420
gaaactccat ctctactaaa aatacaaaat tagctgggca tggtggggca tgcctataat 33480
cccagctact caggaggctg aggcaggaga atggcgtgaa cccgagaggc agagcttaca 33540
gtgagcccag atagcaccac tgcagtccag cgtgggcaac agtgcgagac tccatctcaa 33600
aaaaaaaata aataaataaa agaaaaataa atttatgatc tatttcaaaa ataacacatg 33660
tactttgaaa cagcagagac acatatgaca cggagaatga aattccccat agcgcacccc 33720
caagagacag ccctggtccc cccgtctttc ccgtggacct ccagcggggc agatgctgag 33780
ccgcctgttg tcgagtggcg tgctatcccg tcctccagct cctctgtggc ttacagacac 33840
ccacctgcag ccctgtcttt gcctcctcta gcgcccacca ccttcttgct gttcagccag 33900
aaatctgcca tcagtcggat gatcccggac gaccagcaca gcccggatct catcctgccc 33960
ctgcatggac tgaggaacgt caaagccatc gactatgacc cactggacaa gttcatctac 34020
tgggtggatg ggcgccagaa catcaagcga gccaaggacg acgggaccca ggcaggtgcc 34080
ctgtgggaag ggtgcggggt gtgcttccca aggcgctcct cttgctggtt tccaggctgc 34140
tgcccctgtc cttagcagag ggaggaaaca gaggatggct ctgggtgaat gatgacttgg 34200
gcttcgatta tgtagtcaca gggtatgacc ctgagatgcg tggaaccccg agactgtgat 34260
tatatgtaga aactgggttt ccccgttgtt taagtagtca tggtggggtc agaccccaca 34320
ggacttttgt cttttcaaga aagaaaatgg tcgtgtgtca tgcaggggta gttggtactg 34380
gttaatccag gtttatcctt tattttgtgg gaactgtaca gtcatttctg ctacaatgct 34440
gtatatgctc ttctgaaaga cacctatgca aaatcgcaca gtaaaaatga cacaactcat 34500
agggaaagcg gggccagggc acagccctca aaatctccat caatgacatg taagaaaaga 34560
gaggaacctg ggaaatagca aagtgccttt tgcacattaa atggttagct atatcccaca 34620
atactgtgca ttcgtaaacg ttaatgctgc aataaatacg gcacttcacc ttgggaagat 34680
ctggagttgg cttatgagtg tggaagggtg tagcgcatga gtttttgtga aacactggaa 34740
ggaggattgt gggaaatcaa atggaaagtt ctcaccccag gcgtggagaa gagtgggtca 34800
tggccccagc agtgagccca gggaggtcag agacggaggt gtgtgtgtgg gtgtgaccct 34860
gcgcagttcc ctgccggctg tagtttttttg cattcgctta atgtttctcg tggaggaaat 34920
tgtgcatgag caaatgtgaa accgtgctgt gctcaaattg tcctaataca tcattgcatt 34980
ggaacagatt ggcttttttt tttttttttt tttttttttt tttttgagat ggagtctcac 35040
tctgtcacca gcctggagtg cagtggcatg atcttggctc actgcaacct ttgcctccta 35100
tgttcaagtg attttcctgc ctcagcctcc tgagtaactg ggattacagg catgagccac 35160
cgcggccggc cagatttgca tttttgaaac aactgctagg ctgggcgcgg tggctcacac 35220
ctgtaatccc agcactgtgg gaggccgagg caggtggatc acctgaggtc aggggttcga 35280
gaccagcctg gccaacatgg tgaaaccccg tctctactga atatacaaaa atcagctggg 35340
tgtggtggcg ggtgcctgta atcccagcta ctcaggaggc tgaggcagga gaattgcttg 35400
aacccaggag gcagaggttg cggtgagccg agatcacacc attgcactcc agcctgggca 35460
acaagagcaa aactccatct caaaaaataa aaaatagaaa aacaagtgct gtagcggaag 35520

```
tgagcacttt gcggagtcag gcttgtgtgg cctgttccac aaatgatgtg ctcacggtgg    35580
cctcaggccc acctggagtc tgcagcatgg ggcacaacag gttcattagt gtagaattcc    35640
aggacaggcc tggctcctaa gcagccttct tttacaaaaa ctgcagagcc cgcctgtatc    35700
ctagcacttt gggaggccga agtgggtgga tcacgaggtc aggagttcaa gaccagcctg    35760
gccaacatgg tgaaacccca tctctactaa atatacgaaa attagctggg tgtggtggca    35820
cgcgcctgta gtcccagcta ctcgggaggc tgaggcagaa ttgcttgaac ctgggaggtg    35880
gaggttgcag ggatctgaga ccatgtcatt gcactccagc ctgggcaaca gagcgagacg    35940
ccatctcaaa aaaaaaaaac ctacagagcc acacggcctc tttctccacc gagtgttggt    36000
gtgggagctt gtgttattgt ggtgaaatct tggtactttc ttgaggcaga gagaggctga    36060
gcgcctggag agactttcac atgggtcgcc atgtccgccg tcggtttcgc tgttgtgctc    36120
cccatctgaa ggctggtgcc gtccagacag gctggacgcc cctttccacc agatccttcc    36180
tcccgcagca gtttctagtt acgttgtact gtgaggtctg tgtccttggt tgatggcaaa    36240
agtcagccga attgaaattc agagccatgc ctggctccct ggagcttctc tcctgggcag    36300
ctgtgatcat tgcctctgct gtggtgtggg tggtggaaat ggattccttt catcttgctt    36360
gctacaggtg actgtcacgt ggagtccttt ggagagaggg acgtgttaat tgatggatgt    36420
ggctcccatg ctgagaaagc tcctgggcgt acattgcctt agagtttcat tggagctgcg    36480
ttcttttatg gtgtctgcta ggcagaagtg atgaagactt ggaagaaaac ccagaaggtt    36540
ttccacttaa tttggaaaat gtgcttttcc cctcctgtgt cttttgctaa ggtccagcct    36600
cctgcagcct ccccgctctg tggactctgg ctttgattct ttattaggag tccccctgct    36660
cccccaaaag atggtgtcta aattatcatc caattggccg aggttttgtt ttctattaat    36720
tgtttttatt ttttattgtg gtaaatttat ataacataaa atttgccatt ttaattgttt    36780
tgttattgtt gtttttgaga cagggtctca ccccagtgcc caggctggag tgcagtggtg    36840
cgatcatggc tcactgcagc ctcagcctcc agggctccag tgatcctctc acctcagcct    36900
ctctagtagc cgggactaca ggcatacact accacatctg gctgattttt tgtatttttt    36960
ttttattgta gagacccgct atgttgccca ggctggtctc aactcctgga ctcaagccat    37020
cctcccacct caccctccca aagtgctggg attacaggca tgagccacaa cacccagcca    37080
ttttaatttt tttttttttt tttgagatgg agtctcactc tatcgcccag gctggagtgc    37140
agtggcgtgg tatcaactca ctgcaacctc tgcctcccag gttcaagcga ctctcctgcc    37200
tcagcctcct cccgagtagc tgggattaca ggtgcccatc actatgcctg gctaattttt    37260
gtatttttta gcagagacgg ggtttcacca tgttggccag gctggtcttg aactcctaac    37320
ctggtgatcc gcccgcctcg gcctcccaaa atgctgagat tacaggtgtg agccaccgtg    37380
cccggccttt ttttgttttt gagacagggt cttgccctgt cacccagact ggagtgcaat    37440
ggtgggctct tggctcactg cagcctccgc ctcccaggct caagttgtgc acctccacac    37500
ctggctaact gtattttatg tagagacaga tttcaccatg ttgcccaggc tgggcttgaa    37560
atggactcaa gcagtccacc cacctcagcc tcccaaagtg ctgagattac aggcgcgagc    37620
caccgcaccc agcccatttt acctattctg cagttgacag ttcagtggca ttcagtcagt    37680
tcacgaggta accatcactg ccattcatct ccagactact tcaccttctc ggcagatgtc    37740
cgaaactgtc cgcattgaac acactcctca tctccctctg acagccacca ttctactttg    37800
tatctctctc tgccttctct aggtacctca tgtaagtgga attataccaa tatttgccct    37860
tgtgtgactg gcttctttca tgtgacatgg tgtcctcaag gttcatctgt gttatagcct    37920
gtgtcagaat ttccttcctt aaagcctgaa taataacccg ttgtaaaggc tgggcgcggt    37980
ggctcacacc ctctaatccc agcatttggg gagtccgagg tgggcagatc acttgaggtc    38040
aggagtttga gaccagcctg gccaacatag tgaaaccctg gctctactaa aagtacaaaa    38100
ttagctgggt gtggtggcgc gcacctgtaa tcccagttac tcaggaggct gaggcaggag    38160
aatcgcttgt acccgggagg cagaggttgc agtgaaccaa gattgtgcct ctgcagtcca    38220
gcctgggtaa cagagtgaga cttcctgtct caaaaaaaaa aaaaatcatc ggatggatgg    38280
acggaccact tcttgttatt tatccatcca cgggtgctag gtttcttcca cctttggttg    38340
tcgtgaataa ggccactatg aacatttcct tccgtggtga aggttttgta ctagtgagga    38400
aaaggcgtgt ttgtggtgtt gcataggatt ctggtaagaa agtttgcact aaccataagt    38460
atttgtacta cattaaaatg aaagctcagg ggccgggcgc ggtggctcac gcctgtaatc    38520
ccagcacttt gggaggccag ggcgggcgga tcatgaggtc aggagatcaa gaccatcctg    38580
gccaacatgg tgaaaccccg tctctactaa aaataccaaa aaactagcca ggtgtggtgg    38640
cgggcacctg tagtcccagc tacttgggag gctgaggcag gagaatggcg tgaacccggg    38700
aggcggagct tgcggtgagc cgagatcgct tcactgcact cgagcctggg caacagagca    38760
agactccgtc tcacgcaaaa ctctgtctca cgcaagactc cgtctcaaaa aaaaaaagag    38820
ttcagggttt atgaaactgg ccagccgcgt aaagtttgct gtgttgtttt tgtgcccggg    38880
aggagtgtgg ccagggtgtc acgtcacaca gtacacgttt ctcagatggt ggttctccag    38940
```

actgctgtcc caaagtctgt ttttgcatct ggttcccaca gacccaccct ccacggtgag 39000
cctgattttg gccagggtag ctggaatctt gcttgtcttt cagcccggca gctgtaccag 39060
tccagggtcc acagctagtg gcttttagga aggaatttgt tcagttggct ttgacacatg 39120
gccccctagg gtccacagct ctgtagtgat gtggatgttg ttatctacaa agacacatga 39180
tccttcgtgt ccagatgaaa gtgatgatgt ctttgcagct gcccagcaag gctgtgtgtg 39240
tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg tggtgtgtgt gtggtgtgtg tgtgtgtatg 39300
ggggagggag gcaccctttc catctggggg tgtgtgtgtg tggggtgtgt gtgtgtgtgt 39360
gcgcgtgtgt gtggtgtgtg gtgtgtgtgt gtgtatgggg gaggcaccct ttccatctgg 39420
gtccaagaga ctgggcctgg ggaagacgct tctttttatc tacttagaga ctttgtttta 39480
tttgtatttt tttgagacag ggtctcactc tgtcacccag gctggggtat ggtgatatga 39540
gcatagctca ctgcagcctc ggcctcccag gctgaagcga tcctcccacc tcagccttct 39600
gaatagctgg gactgtaggc gtgcgtcacc atactgagct attgtttttt ttgtttggtt 39660
ggtttaattt tttttgatac agatggagtc ttgctatgtt gcccagacta gtctcaaact 39720
cctgaactca agtgattctc ccacctcagt ttcccgacat tctgggatca caggtgtgag 39780
ccactgctgt ctccctgttt tattaactgc tgaaagacct agataaagaa agtctgaaaa 39840
gacttactat cagagcacca tcctaagatg attccctctg actcaatgga gagggagggg 39900
agcttttcct tcaggcctgg gtggcaggag cccaggtgct ccaggcccca tttgccccag 39960
gccaaatcac tcgggaactt ggatgcagct gtctttcagg gtaacccaaa ggaaccagat 40020
ccccgcaggc agtaggcttc tgggctgtcc tctcctccta cgtcagctca gtaagagccc 40080
ttcgaaggga tgctgtgtcg gaggccccaa aagcccaggc tcatccctga gatgcacagg 40140
gtgggctggg cttaggcagc gctcgagcat ctcctggacg gtgaccccag agagtgtgga 40200
gacggagagt ccttgagagt cactgagaga cgtggctgcc ctgccttccc aagaggggct 40260
ctgagtcatt ccccacactc acctgcccct acccaccctc acctggcccc cagcctcacc 40320
tacccccaca tctgtaccga tccctttacc cgcaccttcc ctacccaccc tcacctcccc 40380
tgtaccttca cctcccccac tcacccgccc ctgcaccctc acctgtcccc caccttcacc 40440
taacccccac cctcacctgc cctcccctca cctggcctcc ttccgttggg gaaggggttg 40500
taagggggcgg cccccaaact gtctgtcctg gtgccctgca gagaaaacag tacgtgaggg 40560
ccgcagtcca aaagcttgag tcctggaagg tggaggagac agggatgtgt tgggaagggc 40620
cccatggtct tggatccctt ctcgactgtc aatggggcct tcatgggagc gccagtctag 40680
tgatgcacag ctgggtgccc ggcgggtggc tgaggaggcc taaagtccga ggcggcaaga 40740
gctcttccag aggctgttgt cctaatcgct ctggcatact caggcgggca cgtagttagg 40800
agctgattgg agaggagaga cccccacacc aatactggga tttgactttc aggctaaact 40860
tgagaagtgt ggcctctgct gtcctgccag agctctccag ccagtgccca gggctctcca 40920
gccagtgccc gggggtctcc accagtgccc gggggtctcc gccagtgcca ggggtctccg 40980
ccagtgccca ggggtctccg ccagtgctca ggagtcttgg tttctttgtc ttacagccct 41040
ttgttttgac ctctctgagc caaggccaaa acccagacag gcagccccac gacctcagca 41100
tcgacatcta cagccggaca ctgttctgga cgtgcgaggc caccaatacc atcaacgtcc 41160
acaggctgag cggggaagcc atgggggtgg tgctgcgtgg ggaccgcgac aagcccaggg 41220
ccatcgtcgt caacgcggag cgagggtagg aggccaacgg gtgggtgggg gtgctgcccg 41280
tccaggcgtg cccgccgtgt cttatgccga atgccagcct ctcacaggct ggggagactt 41340
tccacctggg gatccaatgg gtggctttcc agggtcccaa aagcaaacac aggtttttca 41400
cagcccgtcc gggaaagcag aaagccccaa ggggctggaa ggggaaaggg ggagctctgc 41460
tgagaggtta caaggcagcg ctggccgacg ggagttgcag ttgataggtt ttgtatcatc 41520
cttgttaaac ttgaaccctg tgcagaaatc ccttccacgg catgggggct gcctgttgac 41580
tcgctcctgt tccaccacag ggagctcctg ggcttcttcc tcccagaggc ccccgacgct 41640
cccacctgtt ggtcgtcaga gcttctggtt ggtgggaagg cacccaggac cttgaggtct 41700
ccagagagaa aagccaggga aagagggaga ccgaaaccca tgtgacatga aactcaggct 41760
ccaaactgag cacgggaacg tttggggaca ggagcgcgat ggccttcctc agatagctgg 41820
ggggctggca tgaagacggg agctacagcc agcacaggtc ctgggccggg agcccagaga 41880
ttgagccctg actctgtcac ttactggcca cgtgaccttg ggcgggtggc atagcctctt 41940
ggagactcag tttcctcatt ggtaggagtg acggccacag tggtgcggcc tctgcagcac 42000
acggggggct cggtgggcgg aagccccggg tctataaggc ggctgtgcag gagccagccg 42060
agctggtctc ccaacagcca gggctccggg gtccttagca gctgtggggg gcctgcacct 42120
gtttcccatg gctgctgtca gaaattacca gaagccaggt ggctgagagt aatggacact 42180
tgttctctca cagttcctga gggctgaagc ccgagatcga ggtgtgggca gggccctgcg 42240
ccctctgaag gctctgaggg aacctttggg cttctggtgg ctccaggcac cccttgactt 42300
gtggtcctgt cactccagtc tctctgtctg gctgcacatg gcgtggcctc ttctgtacca 42360

```
ttgaaggaca cttcagttgg atttagggcc taccctcacc cattgtggtc gtatcttgat     42420
ccttcatgac atttgtaaag accctgcttc caaataagct cacattctga ggttctgggg     42480
tgagcgggaa tttggagagc attgttcaac tagtatagaa tgtgacctgt cagcctcggg     42540
cagccctgag aggcaggggc tttccacagc ccagctgggt gccctgggct ccgtgctgtc     42600
cgaggagacg ccatccccac acccgtcctt cacccgccac cctcccgcag gtacctgtac     42660
ttcaccaaca tgcaggaccg ggcagccaag atcgaacgcg cagccctgga cggcaccgag     42720
cgcgaggtcc tcttcaccac cggcctcatc cgccctgtgg ccctggtggt agacaacaca     42780
ctgggcaagc tgttctgggt ggacgcggac ctgaagcgca ttgagagctg tgacctgtca     42840
ggtacgcgcc ccggggcctg ccctaaccgc agacacccgg ccttcattgt cagtaatggc     42900
agcagctgcc acattgtccg agacctgccg tgagcccagt gccgcgccag gggctttgtg     42960
tgtagcgtgt tttgtcctca cactgacagc tgtaggctgg ggttctgagt gagccccaca     43020
gggcagaggc agaaaatgag tctcagagag ggtgagcgag ctgcttgggg ccccacagca     43080
ggagatggag caggactgca gcctagcctc tgcccccagc acctgcgcaa gaagctgctc     43140
tgctctggac tgtgttaggc tgcgagggct ggagagaaat gagagttggt gcttagagag     43200
ggggcgcagg tccccatggc ttttcctctt atgatgaggt agatgggtga agggaggggc     43260
catgcttgca ggggccagtg accgaggccc gccgttggaa ctgatggcct tcatcccgag     43320
cccagcccag gtgggagcag ggctttccga gggcttgtct tgggtcggcc tgcttccagg     43380
gactctgctg cagctcccac ccctgtccaa agcatggaat cccccaggct ccctggcagt     43440
cctgtcaacc tctgtcctcc caagctgagt gtggggcaag ttctggaggt cagcactgct     43500
caggggggcc cacgggctgc ttgcaggggc caaccgcctg accctggagg acgccaacat     43560
cgtgcagcct ctgggcctga ccatccttgg caagcatctc tactggatcg accgccagca     43620
gcagatgatc gagcgtgtgg agaagaccac cggggacaag cggactcgca tccagggccg     43680
tgtcgcccac ctcactggca tccatgcagt ggaggaagtc agcctggagg agttctgtac     43740
gtgggggctg gcagtggggt gggcagggtg gcctctaaac ccgacccctg gaggaggctg     43800
gaggccagtg caagatcctg tgtggcctca gccaggcggt ggtctctgcc agatgccaac     43860
tgttgcccgc tggggttcag cgacatgtcc gaatgtcccg aggcctctga ggttgttttc     43920
ttttgccgca gaacaaatca ccacgaacag cgttttaaga caacaccaac tctttttttt     43980
tttttttttt tgagtcagga tcttgctctg ttgcccaggc tggggtgccc tggtgcaaac     44040
acagttcact gcagcctcga cctctgggct taattaagtg aacaccttgc ctcagcctcc     44100
caggtagctg ggactacagg tgggcaccac cacacctggc taattttttt ttgtagagac     44160
ggggtttccc catgttgccc aggctggtct gcaactcctg ggcacaagct atctgcctgc     44220
tgtggcctcc caaagtgcta ggattatagg tgtgagccac tggcctgaca acacccacgg     44280
attgtctctc agttctgtaa ggcaaagtcc aggcacagcg tggctcacct gggttctctg     44340
ctcagggtct cacggggcca gaatcaaggt gtcaggaacg ctgggccctc agcggaggct     44400
ctgtggagaa attagcttcc ttgctcactc agcaggtagc agttgtggga tcgaggttct     44460
gttttctctc tggttattgg tcggggacca ctctcagctc ctagaggcca ccacaggtcc     44520
ttgccccgtg gccctctctg cctcagcagt gggggctccc tgcgtcagtc cctcccacac     44580
cttgagtctc tctgatttgc ttctaaaggg ccctgtgatt cggctcagcc acctttagat     44640
taggttagcc tccccctttga tagactccaa gtcggctgat taataacctt aatcacatct     44700
gcagaatccc ttctgccaca taaggtcatg acgccgtgct ggggactggg gtgggaaatt     44760
acggggtcat ttaggattct gcctgccact gccttgctgt gtcccagggc ttgggggagg     44820
ggcctccaca gctgggacca cagtccttcc tcccctccat ggtaaccatc tgaggattac     44880
ttgagaccag cctgggcaac atggtgagaa cccatcccta caaaaaatac aaacaaaaag     44940
ggaccaggct gggcttggtg gctcatgcct ataatcccag cactttggga gaccaaggtg     45000
ggctgatcac ttgaggttgg gagttcgaga ccagcctgcc caacatagtg aaatcccgtc     45060
tctactaaaa atacaaaaat tagctgggtg tggtggcagg cgcctgtatt cccagctact     45120
ggggaggctg aggtgggaga attacttgaa cctgggaggc ggaagttgca gtgagccaaa     45180
attacgccac tgcactccag cctaggcaat agagtgagac tccgtctcaa aaaaaaaaaa     45240
gggccagggg tggtagtgac aaagagaccc tatcccaaaa aaaccgaaca ctgaatcctt     45300
gagactgagt aaggacactg tgaaattttt ctgggtgggg cagggaacag agcgtcttct     45360
gtcatttctt ccacctgggt gtggtcagct ctccctccaa gctgcctcct cttcttctca     45420
ttgtccgggt gttggacaca tttggttaac tggatagaat aacgcgagtt cccagggact     45480
tggtccattt gctattttat tttattttat tttattttat tttatttatt tatttattta     45540
tttatttatt tattgagatg gagtttcgtt tttgtcgccc aggctggagt gcagtggcgc     45600
gatctcggtt cactgcaacc tctgcctccc aggttcaagt gattctccta cctcagcctt     45660
ccaagtaact gggattacag gcacccacca ccataccagg ctaatttttt tgtattttta     45720
gtagagacgg gttttcgcca ttttgcccag gctggtcttc aactcctagc ctcaggtgat     45780
```

```
ccacgcacct cggcctccca aagtgctggg attacaggca tgagccacca cgcctggcac     45840
catttgctat tttaattccc atgtgtatta gtgtcccacg gctgctgtaa caaatgacca     45900
caaactggat ggcttaaagc aacagaaatg gattccccca atgtgctgga gaccagaagc     45960
ctgcgaccaa actgttggga gggctgtgct tcctctgggg gctccaggga ggatctattt     46020
gttggccctt ccagtgctgt gggtgccagc gttccacact tgtggatgcg ccgcctcaac     46080
ctctgcccat cttcatgtgt ccatctcctt tgtgtctgcg tctttacctc ttcttcttgt     46140
ctgtgttgcc tcttataagg acgtttgtca ttgggtttag ggcccaccca aatcatccga     46200
gatgacctcg tcttgagatc cttaacctgc aaagaccctt tttccaaaaa aaggttatgc     46260
tcacagattc taggccttaa gacatgggtg tatctttctg gggggcacta tccaacccct     46320
tatacaatga aagacgggaa gagggccagg tgtggtagtt cacgcctgta atctcagcac     46380
tttaggaagc tgaagcggga ggatcacttg agcccaggag tttacaagta gctaggcaac     46440
atgatgagac cccatttcta caaaaagtga aaaaaaaaaa aaaaaaaaaa aagccaggtg     46500
tggtggctca cacctgtaat cccagcactt tgggaggctg aggcaggcag atcacgaggt     46560
caggagattg agaccatcct ggctaacacg gtgaaacccc gtctctacta aaaatacaaa     46620
aaattatggc cgggcgcagt ggctcccgcc tgtaatccca gcactttggg aggccgaggt     46680
gggtgaatta caaggtcaag agatcgagac catcttggct aacacggtga aaccccatca     46740
agatcacaag gtcaagagat ggagaccatc ctggctaaca cggtgaaacc ccgtctctac     46800
taaaaataca aaaaattagc cgggcatggt agcgggcgcc tgtagtccca gctgctcggg     46860
aggctgaggc aggagaatgg cgtgaacccg ggaggcggag cttgcggtga gccgagatcg     46920
ctccatgcca ctgcactcca gcctgggtga cagagtgaga ctccgtctca aaaaaaaaaa     46980
aaaaaaaaaa aaaaaaagaa aattagccag gcacagtggc aggtgcctat tgtcccagct     47040
acttgggagg ctaaggcagg agaatggcat gaacccggga ggtggagttt gcagtgagcc     47100
gagatcatgc cactgcgctc cagcctgggc gatagagcaa gactctgtct caaaaaaaaa     47160
agccaggcat ggtggtgcat gcctgtagtc ccagctactc aagaggctga ggcaggaggg     47220
ttgttcgacc cacggagatc aaggctacag tgagccatga tcgcaccact gccctccagc     47280
ctgggtgaca gagtgtgacc ctgtctcaaa gtaagtaaat aggaggagag acaagtgggc     47340
agttcagact gatggtatgg gcacagtaga gactggtgca gacaggctgg cctgtgatgt     47400
caagcaactt ctgtaattgt ttccggcatc catttgtgtg tcaatttccg tgtcagtagg     47460
aagactctgt aggctgccaa gaggaataag tgggaggatc ctcccagaga ggccgggcct     47520
gcaggagggc cagttctcat gagttctcat ttggccccta ccctccaggc tgtggttctg     47580
aggtgggaga cagagcctga cctctgtttg tcttgttttg tctttgcagc agcccaccca     47640
tgtgcccgtg acaatggtgg ctgctcccac atctgtattg ccaagggtga tgggacacca     47700
cggtgctcat gcccagtcca cctcgtgctc ctgcagaacc tgctgacctg tggaggtagg     47760
tgtgacctag gtgctccttt ggggtgatgg acaggtacct gattctctgc ctgctaggct     47820
gctgcctggc atccttttaa aatcacagtc cctgtggcat ccagtttcca aagctgattg     47880
tgtcttcctt tgccctcctt tcttttctac tatgtgcatt cggtgctatg aattttcctc     47940
taagtactgc gtttcctgca tctcacaaat tttgttacat tttcattttc aggtagtttg     48000
aatattttta cacttctcct gagatgacat ctttggctca tgtgttattt agaagtgttg     48060
cttagtttct aaagagttgg ggcttttcca gctgtctctc tgcaactgat ttctaattta     48120
attctactgt agtctgagag cttattttat atgatttctg ttattttaaa tgtgttgggt     48180
gtggtgtttt tgttgttatt gtttttgtgt ctttttgttt tgttttgctt cgtttgtttt     48240
gtttttgaga cagtgtcttg ctctgtcact caggctggag tgcaatggcg cgatctcagc     48300
tcaccgcaac ctctgcctcc cgggttcaag tgatcctctt gcctcagcct cctgagtagc     48360
tgggattaca ggtgcacgcc accataccca gctaattttt gtatttttag tagagacggg     48420
gtttcaccat gttggtcagg ctggtctcga actcctgacc tcgtgatccg cccacctcgg     48480
cctcccaaag tgctgggatt ataggcgtga gccactgtgc ctggccatta ggtgtgtttt     48540
atcacccagc atcatgcagt ttatcttggt gaatgttctg tgtactcttg aaaagaatgt     48600
ggattctgct gttgttgggt ggagtgttcc agaaacatca attagatcca gttggttaat     48660
agtgctcatc aggttgtctc tatccttcct tcctgactgc ctgcttgagc tgtcagttat     48720
tgacaggggt gtggagtctc caactctaat ggtggatttg tttatttctc ctagtagttc     48780
tatcttttc tctccttcta cccttgatcc tcttctcccc ctagggcttc ctggtgttag     48840
tggtgggaga gtggggtagt gaagaacctg gactttaggg ccaaagaggc cagggttcaa     48900
atcctggctc tgtcacttcc cagttgagtg accctggctg gtgcctgaat ctctgtgagc     48960
ctccacttcc tcctctgtga aattgagagc acttacctgg caggctgtca tgggcatcaa     49020
gtaacagggc actccacctg gaccctgaca cgtgatgcac aggaatgcca gctgctatgc     49080
catgggtgtg gcagtagtaa taaagtgacc atctgtatcc tcaccacagt gaagcctgtc     49140
cagggctttc tctcctatgc cccatgcct ccaggtggcc ttggatcctg ttggttctgt     49200
```

```
gctctgctca gcgacctttc tcccgtggga gttcctgggg gttcagcttc atcctacaga     49260
cagcagcaca cactggctgt gcaccctttt tttttttttt tttttttttt tgagatggag     49320
tctcgctttt ttcgcgcagg ctgaagtgca gtggtgtgat cttggctcac tgcaacctct     49380
acctcctggg ttcaagtgat tttcctgcct caccctccca agtagctggg attacaggct     49440
cccaccacca cgcccggcta atttttgtat tttcagtaga gatggtgttt caccatgttg     49500
gccaggatgg tcttgaactc ctgacctcag gtgatccgcc cacctcagcc tcccaaagtg     49560
cagggattac aggcgtgagc caccacaccc ggagtgccgg ttgtttttag cagtttgtct     49620
tgttcctgga gagactggct cctgcccagg agctcgggga gtagggccgc ggggtgctgc     49680
ctcacacctc gagtttggcc gtaagcagag gggacatttt gtgactgtcc ccctcctgag     49740
cttcccagca gcttttctcc aagttacagc ccaaaagctc aggtggattt gcaacccaac     49800
ggtgtctgtg cacctcccac tgatgcccga actgccctgg ccaagaaacg gggccgtcag     49860
aacgctgcac taactgcagc cttgggcctc catgccagag gccatgccct tccatccacc     49920
accccctggc ctgggccctg ggccctcctg gctcgggaac tccaggcccc ttcctcacgg     49980
ctcgagagac gtgtatttac cgcacaggtg cttgtcattc tcttgtggcc tcttctccag     50040
ggagatcaca gaaggacagg gcctcactga ggtctcggac atggaccctt tgatagtggc     50100
aggagccagg ctgggcaaga ggcggccaca gtcacctcag cagtgccatc accaccgcca     50160
ttcagccctt ccctgagccg ggcgcgcccc tggctctggc cccagtgtcc cagttacagc     50220
tcacaggagc ttgtggtgcc cagcggctgc ttctgattga gagtcgaggt cggaggcttt     50280
gggaggctga gaggctgctc ggtttcacaa ctgctgaggg agacttgggc tccatctcag     50340
gtatgcccca tgtcgccctc aacctccagc caccggtcct ccgtgtcccc catggccagg     50400
cacggcttgc agacatctgt cgttggctcc tctcagccgt cgtgggctga ccctggcacg     50460
tcctcctgtg gctgagccca gtggggacag ctgcttcctt ttattaccct agaactctcg     50520
tctttgatca ggccccctcc cctatgccac acagtccctg tcactcgggt gagcccagta     50580
gtcatgggga aggcctgcgg gttccaaaca tccaaaggct tgcgtgcagc atgacagctt     50640
gaaaccgatg tttttacct tgatcagatt tcagcttggc gggggctttg ctcagctttc     50700
agtgaggcct gggccgattt cccagcatcc cctcctgagg ccagcctctg tttcctgtga     50760
ttttctgcac aaagtgggag ggaggagtcc taggaaatgg ggggccacct cgaagcctag     50820
gcctcctctg gcttctctgt gccagtgccc ccacgctttg tgtctgtgtc cccagcccat     50880
gggactctgc tattccctga gtgctgccgc atgcccagcc cgcactgagg acgtggagcc     50940
ccgaggggca ggatggcctc catggtcaca cgtaggaagt ggcctccacc ctccgatgat     51000
cctctccctc ctccctttca gcgccctccc cgggggtgtc ctcagccctc ctgcctgtgc     51060
tttgtcccgt cttctgcagg cgcctgggac gtgctgacag gtcctctgcc ggctcctgcc     51120
ttgctatgcg cacgctggtc accacagagg cctggccctt cttctgtagc agtcccacac     51180
ccgcaacagg tgtggctgct gaccacctgc tttctgcccc tctggtcctg aggagggcgc     51240
agtgggcact caggcgtggc tgagcagatg tgtgttgccg ggaggaggaa ggactgctcc     51300
agtcagggct gaatttccca cccggagcat ttctgctgta tttggtgtag cgcctgctgc     51360
ttaaagctct gattcccagt tggcaccctt tcccttctgc attgaaaaac atacggatgc     51420
atgtcttctt gcagtgaatg tgtattctcc cagcctctct tctgggttgg ggctggaggt     51480
ggagcggcac acaggagccg cagcgatgga ggatgtgcgg gtgcagcacc ccgtacagca     51540
gggatgccaa acccgcgctg agtccctctc aacttctgct ttgaagccca gtcacgccat     51600
tgcctgggtt ttgctgggcg gggctgcgtg tgatgttctc ctctgtccct cccccagagc     51660
cgcccacctg ctccccggac cagtttgcat gtgccacagg ggagatcgac tgtatccccg     51720
gggcctggcg ctgtgacggc tttcccgagt gcgatgacca gagcgacgag gagggctgcc     51780
ccgtgtgctc cgccgcccag ttcccctgcg cgcggggtca gtgtgtggac ctgcgcctgc     51840
gctgcgacgg cgaggcagac tgtcaggacc gctcagacga ggcggactgt gacggtgagg     51900
ccctccccgt caaggctctg ccaagaccct ggccctgccc tccgggatac gagcttgggg     51960
ctgcctccgg cctcacagga gtagggggctc tgaaaacctt tgcttgcagg gagattgcca     52020
agtctgtctt ttaggcccaa caaggaaaac tctgcagttc cacccatcct gtcccaccag     52080
gtagtgtggc ttgaaggcag actgtgaggg tctatctcac cttcctgcat taggtcagga     52140
gtttcacaga aacctgaggc acattcaggg gtgggctgca gaggtccatg gctcacaccc     52200
tggaaaatcc gcccccaaaa gacagtgctg tctccactga ccagtctgtg ggatagtgct     52260
taagcctgag tggtttctat caacatgtag aatcaggagg tataaagaga tttgctcagg     52320
catcctgggc cctctctgac cagcaggatc ttcctttaga tcttgacagt gaaacacatc     52380
tcttctgtgc cccctgtgag ttttctttca ttcattcatt cattcattca ttcattcatt     52440
cattcattcg agacagagtc ttgctctgtc acccaggctg gagtgccctg gtgtaatctc     52500
ggctcactgc aacctctgcc tccagggttc aatcgattct cctgcctcag cctcccgagt     52560
agctgggatg acaggtgcgc accaccatgc ctggctaatt tttgtatttt tagtagagac     52620
```

agggtttcac catgttggcc aggctggtct cgaactcctg acctcaggtg atccgcccgc 52680
ctcagcctcc caaagtgctg ggattacagg catgagccac cgcgcccggc ctgagttttc 52740
cttttatgaa ggacctgctt ggttggttgc ctgccacatg ttgtcagcac catgggccca 52800
ggactgctga ggagctgttg atgccctcgc tctcccagag ccaccggctc tgttagataa 52860
ttcacatgca gtctggccac tgtcctacgt cctcattcac aaagagcaga catttcgtag 52920
aagatgaggg cctgggagta acctccctgc atgtttttct ataaaggcat agtggttaag 52980
tccttccagc tcattgacca ttggagaatt ttatggaggc tgtagactag ggctggtaa 53040
actaagggcc caggggccaa atccagcctg ccacctactt ttgtaaataa agttttcttg 53100
gtgcacagcc atgcccattc attcatttgc acaatgtctg tggctgcttt catgccaaaa 53160
gcaagagaac tgagtggtta tgctggagac ctacggcctt caaagcccca gacctcacgt 53220
ctggcccttg acagacagag cttccccagc cctgctgcgc atcctggccc agcatgtgct 53280
gtgtgtgtga tttcagcttg caggagccgt ggttaggaat tgtccctgtg ttggtccatt 53340
ttgcattgct atgaaggagc acctgaggcc gggtagatta tgaaggaaag aggtctgtct 53400
ggctcatggt tctgtaggca gcaccagtat ggcacccgca tctgctcagc ttctagtgag 53460
gtctcaggaa gctttgactc atggtgaaag tcgaagcggg agcaggtgca tcacatggtg 53520
agagagggag caacggagag agagagagag cgcctctccc tcttgccctc accttgagag 53580
gagatgccag gctcctttaa gtaaccagct cccatgtgaa ctcacagtga gagcccattt 53640
gctactgcgg agagggcacc aggcatctgc tcccatgacc caaacactgc ccaccaggcc 53700
ctacctccaa ccttggggtc atattttatt ctgttctatg ctatgctatg ctatgccatg 53760
ccatgccatg ccatgctatt cctattctat tatttgagac agaatctcgc tctgttgccc 53820
aggctggagt gcagtggcat gatcttggct cactgcaacc tccacctccc aggttcaagc 53880
gattctcctg tatcagcctc ccgagtagct gggattacag gcacacacca ccacacccgg 53940
ctaatttttg tatttttcaat agagatgggg tttcaccatg ttggccaggc tggtctcaaa 54000
ctcctggcct caagtgatcc acctacctcg gcctcccaaa gtgccatgat tacagatgtg 54060
agtcactgcg cccagtgagg gtcacatttc cgttgagatt tggaggggca gacgttggag 54120
ccatctgagc cccctcgtcc cgctctagct tctcctcccg tgtgccccgc ggtgctggtg 54180
gcaggccctt acgccggttc tggctgcatg ctctgttcca gaagctttct tccctgcttg 54240
gttaccagaa aatcatccca tccattacaa ggacagggtc cccttatctc ccattcccag 54300
ggcaggacac cgggggcagg gcaggtgggg aactgagcaa gttctctggg ggcaggcgtg 54360
gctatggctc cctctgggtg ggcgtctggg gaggggtgga ggcagccgtc agcgccctgg 54420
cttgctcttc ctccctggcc agagactgtg gccttgtgct gctcccgtgt gggctgcctg 54480
cacctccagt gggttgtgct ccctccccctc ccctccccctc aagctctgct gagcaccact 54540
gccttccaca gcccccactc tcgggaggcg aggctcctcg tggccattcc tgtccttggc 54600
acccaccccc ccaccaacct ggtagagcct tgggcggggt ctgttactcc ttgcatggcg 54660
tagacctccc cacagtaggc acctgacaca tacctcctgg ggggcaggca ggaggtgcgt 54720
tgaggtctca gccctggcag tccctcccct gcgtggcata ggcctcgcca cagggtcatc 54780
gagggtgggt ggagactgta ctagaccact ccccgctggt cctagaaagg gtcccatctg 54840
tctgctctct gtttggagtc cagaccttgg ttgctgtgcc ctgcatggtg ggctgggggg 54900
caccctccag cctctctgag tgcatggcct ctccttgcag ccatctgcct gcccaaccag 54960
ttccggtgtg cgagcggcca gtgtgtcctc atcaaacagc agtgcgactc cttccccgac 55020
tgtatcgacg gctccgacga gctcatgtgt ggtgagccag cttctggcac ggggaagggg 55080
cgtccgggct gggttccccc aggaacgtgg agtttagggg aggagacgtg cctttccagc 55140
ggggctgggg gctgtgtggg agactcaggc ggctgggagg ctccttgcgg gaggcaggga 55200
agcctttccc agggcagcgg ccaggaggac agactgtgag ctgtgggctc ggcggctaca 55260
gagtctgcct cagtgggcgg ggctgatggt gtccaggtgc ctgcagcacg cacccaccca 55320
cgggaccttg ctgagcagcg tctgtcaggc agcaagatta cccgagggct gcagtggtcc 55380
tgttccctgg cagcttactg tctggctgag gaggagtgat gttcacatat gcacacatgt 55440
catgtgcaca cacatgtaca tgacaacatc ccacatgctc ctcaaatagc atgacctgta 55500
cagtcacgga tatagggcct aggggatagg aggccaagac agtcagggaa gactttccag 55560
aggcagtggc tcctgaaagg ctgtctgatt caggcaggaa gggagctgag ttcagatagg 55620
aagtagcaat gagtcattgt gtctggggac atggccactc cttcgctgca gagggacctg 55680
ggctgagagc tcctctctta tggctgcagt cgggagagaa gtctgttggg gggagaaggg 55740
ggcttcctca agggactccc tgtgcccttt ggcaccttcg tgccaggtca ggcttgaggc 55800
ctgaaggcag tggtgggggc caccaagggt cgcctcctct gctgggcaag ttcccagtct 55860
gacgggcctg tgccgtgggc cccagctgtg gggcgctgt tgatgcgcag ccaggcctcg 55920
ccgccagagc ccgcacgctt ccattccgct gacttcatcg acgccctcag gatcgctggg 55980
ccggccctgt gggagagtga atgtggcttt tgccaaagtt gagtctggag cctggaaact 56040

tccctatggg cagccttgat agtggagtgg cccaaggagc ccacccagcc gaccctgccc 56100
ctcccgtggc tggtgggcgg caccaggggc tgcctggctt tgctcgttca ccaacatcac 56160
ctgggctggc cagggcgcgc tcacttctgc caccaccgag ggccctgggc gaaggagtga 56220
ataccaggct gccttggcag ggatgtgttg agggctgtgg ggagtcggac agcggcgggg 56280
gtcagaggag gaggagggtg caccgtgcag gctgaagggc cacgttaccc tgaggttggc 56340
caggctcccc aggcctagcc tcccagctcc cccactttct ccccaccctc caccagtggc 56400
aaagccagcc ccttcagggc gcacggtgtc tgcccccaag gagggcccat tccgttgggg 56460
ttaatgttgg ccacctcttt ctgtttgtct ctggcagaaa tcaccaagcc gccctcagac 56520
gacagcccgg cccacagcag tgccatcggg cccgtcattg gcatcatcct ctctctcttc 56580
gtcatgggtg gtgtctattt tgtgtgccag cgcgtggtgt gccagcgcta tgcgggggcc 56640
aacgggccct tcccgcacga gtatgtcagc gggaccccgc acgtgcccct caatttcata 56700
gccccgggcg gttcccagca tggccccttc acaggtaagg agcctgagat atggaatgat 56760
ctggaggagg caggagagta gtctgggcag ctttggggag tggagcaggg atgtgctacc 56820
ccaggccctc ttgcacatgt ggcagacatt gctaatcgat cacagcattc agcctttccc 56880
actgagcctg tgcttggcat cagaatcctt caacacagag gcctgcatgg ctgtagcaac 56940
ccaccctttg gcactgtagg tgtggagaaa gctccttgga cttgaccttc atattctagt 57000
aggacatgtg ctgtgttgtc cacaaatcct catgtaccct agaaatgaat gtgggggcgg 57060
ctgggctctc tccagagctg aaggaatcac tctgtaccat acagcagctt tgtcttgagt 57120
gcagctggga tttgtggctg agcagttaca attcctacgt ggcccaggca ccaggaacgc 57180
aggctgtgtt tgtagatggc tgggcagccg caccgcagag ctgcaccatg ctggtttgta 57240
tcacatgggt gaccatggta tgtctaagaa ggtggagtcc ctgtgaggtc tgcaggtgcc 57300
cccacagctc caggccacct tgaggattgc ctctgcctgc ccagccctga gttccctctc 57360
ccctgtcctg tcccactgtc accccaagcc ggcctcattg ggagcctgtt ggatggcagg 57420
gtatagatgt aacctgattc tctctgggga gcggggttat ctggcttctc aagagctcct 57480
aggagcccac agtggtggca ccatcacagt cgcagcagcc cccagagaac gcggccctgt 57540
ctgttcctgg cgtgctctgt gctgccccgc ctgggttccc tgccccagtc gcaggcccct 57600
tggaggaggt accatgtgtc tcccgtttca cagatgagcc ccggggagct cactctagta 57660
gtggccagag aggcctgcgg ctcagggagc ggggcacatt tccaacagga cacaccgccc 57720
tggtctgagt ctcgtgggta gtgggagcag aggagagcgc cctatgtctg tggggcggct 57780
tggctgagcc tggaagccac ctgacctccc ccgtcccttc cctgccaggc atcgcatgcg 57840
gaaagtccat gatgagctcc gtgagcctga tggggggccg gggcggggtg cccctctacg 57900
accggaacca cgtcacaggg gcctcgtcca gcagctcgtc cagcacgaag gccacgctgt 57960
acccgccggt gaggggcggg gccggggagg ggcggggcgg gatggggctg tgggcccctc 58020
ccaccgtcag tgctggccac cggaggcttc ccgggttcct gggggctgtg ccaccgcctc 58080
tgaggcatgc ttgctttctt cccttttcaa acccttctgc ttccttcttt aatgacattg 58140
ttgattgtgg ataatctgaa aactacacaa aaatataaag agccaaaatc tcacccaaat 58200
ccacctccta gagtggctgt tgggctccgt cagcatccag gcggccgtct gtgttccgca 58260
cggcccagcc catcgatagc cgcctgcacc aggcctgtct gccctctgtg agcctcccca 58320
cagggttccc tccacaaaca ccctgttctc ccacccaggg ctggctgctt cctggaaaac 58380
agctggatgg ttttgtgcat gacagacaaa cacagggtga ttttcgtggc taaaatactc 58440
cctggagctt ttggcagggt gaggggctgg ctccagctga gccacgcctt gagtgaaatg 58500
actgtgagga gaataaactg ccgctgccct ccaggatcac tggggctggc tggggagaac 58560
ccccgtttct gggagcacag tcccaggatg ccaaggcgag cttggtgccg agatgtgaac 58620
tcctgagtgt aaacagcggg ggctgacttg acatgctttg tatgcttttc atttgttcct 58680
gcagctgtat gcccctaagg tgagtccagc cccctctgc ttcctctggg gcctcgccag 58740
tgagccccac cttgctgggg ctggttcctc ctgcccttct gggtatccct cacatctggg 58800
gtcttgtctt cttgttttct ttttcttttt tttttgagac ggagtttcac ttttgttgcc 58860
caggcttcag tgcaatggtg tgatctctag gctcaccgca acctctgcct cccaggttca 58920
agcagttccc ctgcctcagc ctccctagta gctgggatta caggcatgtg ccaccacgcc 58980
cagctaattt tgtattttta gtagagatgg ggtttctcca tgttggtcag gctgatcttg 59040
aactccctac ctcaggtgat ccgcccacct tggcctccca aagtgctggg attacaggcg 59100
tgagccaccg cacctggcct ttttcttttc ttttcttttc ttttttctga gacagggtct 59160
cgctctgtca cccaggctgg agtgcaatgg tgtcatcatg gctaactgca gcctctacct 59220
tctaggctca agcaatcctc ccatctcagc ccctaagtag ctaggactgc acgcatgcat 59280
ccccatgccc agctaatatt tacatttttt gtagagatga agtttcacta tattgcccag 59340
gctggtctcc aactcctgga ctcgagcgat cctcctgcct cggcctcccc aggtgctggg 59400
attacaggcg tgagccaccg tgcctggcct ggggtattgt cttcttatgg cacctgactg 59460

tggtgggccc tgggaaggaa gtagcagaag agggttcttc ttggtttcct ggacagtaac 59520
tgagtgttct ggaggcccca gggcctggct ttgtttaggg acaaagggaa ctggtaacca 59580
gaagccgaga gtttaaacac ccactgccct tcttccctgc tcctgctgct gcaacccagc 59640
ttaaccagcc aggagtgcta ggaacccaag cagggccccc gagcacacag caggcagctc 59700
acgaattctc ttttcctgtt ctcccttggg agctgggagg atcttaatca ggcaataaga 59760
gatggcactg agcagccagc taattttta aatcacttta ttgtttaacc atatgactca 59820
cccacttaaa aaagggtaca gttcagtggg ttttagtgta ttcacagatg tgtgcaaccc 59880
tcaccacagt taattttaga acattttcct gcccctaaaa gaaactctgc atgaagccag 59940
ctgtttttaa attagcaaag ttattttgca tcctttaaat atatgttcat ggtacaaaat 60000
tcaaaagata cagaagagtc tgcagtccaa agagactccg cccccatgac gccaagcagg 60060
catccctggg aggcatggcc tcctgcagtg tgtttcttct atgtccccc aggggtcatc 60120
tgtacatatg caagcataca agagcgtgga ctttgttttc caagccagaa gataattgta 60180
gatttatgtg cagttgtgag aaagagcaca gacccattta tcctctgcct ggtttccccc 60240
agtgctgcct gccatcttgc atgacttcca ttcctatcat aagcaagaca ctgataacga 60300
ttctttcacc ttattcagat tgacataagt gtttttgtt tgttcttgag acaaacttcc 60360
tctgtcaccc agtgggagtg cagtggcaca atcacagctc actgcagcct caaactcctg 60420
ggctcaagcg attctcctgc ctcagtcccc tcaagtagct cagatggcag gtgtgcacca 60480
tcatgccagg ctaattttta aattttttgt ggaggtgagg cctcactaaa tttcctgggc 60540
tagtcttgaa ctcctgagct aaagtgatcc tcctgcctca gcctcccaaa gtggtaggat 60600
tacaggcatg agccactgcg cctgggctga catatgtgtt ttcgtaagcc cgaaagatag 60660
catctgaaga gtcaacattg agccttgcct tttgctgcta acgatgtata aaagctgctg 60720
ttctgagcat ttcggaggct cccagctgcc gtgtgcaccc tgcctagagc tctaccgtaa 60780
cccatctccg ggaggaggtg ctattgtttt cctcattttg caacaaggag gctgaagaac 60840
tgagcatgaa ccactggcct gggtcgttcg gttggtaggc agtggggcca ggccatccaa 60900
ctcacaacca ccttctactc tgcttccccc gcaccctgaa gtttgttctg ttttgaggac 60960
acagccgtca cattcttggt ggctgaacag cactccttgt caggcgtggc tgggccccca 61020
ctggagggca tcatggtcct ctctcctgct gcggttgaac cttggctgtt tcaaccactc 61080
ctgccaagtg gccctctgaa agggacagtc catcttttct cagcagaggg ccacactggc 61140
aaaacggtcc ctggcaccct ttctctccac ctgtctaata tagagtaaaa atggtatcat 61200
gttaagatct tcatttatat ttattttatc atgaatgatg taagcatcat tttgtgtgtt 61260
taagaacctt tgggcccagc gtgatggctt gcagctgtaa tctcagcact ttaggaggct 61320
gagatgagcg gatcacttga ggccgggagt ttgagaccag cctggccaac atggagaaac 61380
cccgtctcta gtaaaaattt aaaaattagc cgggtatggt gatcccagct acttgggagt 61440
ctgaagcatg agaattgctt gaacatggga ggcggaggtt gcagtgagcc gagatcgcgc 61500
cattgcactc cagcctgggc gacagagcga gactctgtct caaaaaaaaa aaaaaaaaag 61560
aaaagaaaag aaattatcaa tctcctcttt tatggcatat atatatatat atatatatat 61620
atatatatat atatatattt tttttttttg gttatgttca gaaaggcctt ccctgctctg 61680
atcataaaaa acaacttatt ttcacactct ctctcttttt tttttgagac agagttttgc 61740
tcctgttgcc caggctggag tgcagtggcg caatctcagc tcactgtaac ctccgcctcc 61800
cgggttggag tgattctcct gccttacctt cccgagtagc tgggattata ggcatgcacc 61860
accatgcctg gctaattttg tacttttagt agagacgggg gtttctccat gttggtcagg 61920
ctggtctcga actcgcgacc tcaggtgatc cacccacctc ggcctcccaa agtgctggga 61980
ttacagacgt gagccaccat gcccagccca cactctcttt cttaacgtcc tcctcctttc 62040
gttttacgtt cacatcttta attcttctgg gatgtaatta gatttgatga gcaaggtggg 62100
catccagctt gtttcttggc tgatggctta tgggtggcgt gaattagtcg gggtctatca 62160
ggaggcagaa actctatgag aatttgaaca gagaaagttc cgtctacagg cttattacca 62220
gggactggaa tagcagaaat tgaacagtga gatgtacaga gaactctaag aatgcaggaa 62280
taggccaggc atggtggctc acacctgtca tcccagcact ttgggagacc aaggcgggtg 62340
gatcacctga ggtcaggagt tcgagaccag cctggccaac atagtgaaac cccatctcta 62400
ctaaaaatac aaaaaaatta gctgggtgtg gtggcgcatg cctgtaatcc cagctactcg 62460
ggaggctgag gcaggagaat cacttgaacc tgggaggcag aggttgcagt gagccgagat 62520
catgccactg tactccagcc tgggtggaag agcggaactc tgtctgaaaa aaaaaaaaaa 62580
aacaagaagt tcaacttgaa gggaaaaatg ccgtattgtc tttccctttg ttatgtcacc 62640
agggcacagt ccatcccagg ctggcgctga tccacgggct ggagaggggc tgccccagaa 62700
gaggacatgc caggaagggc ttggctggtg ttcaggagcc caggccaggt caggtcaaga 62760
ggtgttgagg ctggacggga gaggccagct aggggctcat gtaggatatg aggggtcggc 62820
ccatttcaac gtggaaactg agctcttctg cttctctttc ttcttcactg cattaagatt 62880

```
caataccgct tgggaagcag gtatttccct tcctataaag gatggttggg agcctgagtg    62940
ttgggagaaa gtgtagccgc tgagttacta acaactaggg ctgccgtcaa gcctatgggg    63000
aaagagagaa gaggacattt ggaaggagag agatcaagct gtggcaccct gggagaggac    63060
cacagaaaag aggccagtga gggggttccc cggtggcatc tgaaggtgtg gcccaaccag    63120
gaggtccaga ggctgccagc cgagtggccc aggagaggga acctcacagg ggctgagtgg    63180
gacccaagcc ctatccaccg tcctaaccac ccacatttct cgggaacaag acctcccaca    63240
gtggcctccc cggcagtgga aatagccaaa ctggcaacat ggactttctt caactgcccg    63300
ggcgatgctg cctcagtgcc ccagggcagg caggaagctc ccacacccat tctggaatga    63360
ggggttggag gaaggctgag ctgagcaaag gacccatctc tgctctggtt ggtggggagg    63420
gagcccatta tacaagagac ccctcagggc tcagtgaggg gtgacagaga cttggggagt    63480
agtggctgtc actgcagagg tgagagggtt tggagagaag gtacatgcct ttttggccac    63540
attgagtagc acctggtagc cagttagtaa cgtgtattgg ataaacaaaa gattaaacgg    63600
atgcaaaaaa aaatgttggc tttgcttctt tttacccaaa cctcagttcc ctcaagtaga    63660
ttctgggaac acccccctacc tggctggact gttgtgaagt ttaaataagc caggttaact    63720
tcacctcctc ctttaagaca cagctcagac actgcctcct ccaagaagcc ccctctggct    63780
tcctgtgtga atatgacggc cctctgggct ctagggtatc ttagaacaat gcttccttat    63840
ggctttggaa ccccgctgtc tcctggattg ggagcaaatg caggggagga gccacacctg    63900
actaatctct gggtctccca gcacataagt ggcataaggg cagggctgtg cccgcttcag    63960
gcacttactg aaggatgtac ttggcagagg gtaggcagcc ggcggatgag cccctcactc    64020
tccccagctg actgcgtggg cgggaaaggc gggttcagga gacccagcct ccctgggctg    64080
tcaccacctc tgcacatcca gccccattga tcaagggttc aatttttggg gtcctgttgg    64140
gaggccagga gactctctcc aggcacttct tccaggtctt tgtgttaggg tgtgtgtgtg    64200
tgtgtgtgtg tgtgtgtgtg tgtgttgttt gttttatttt atttatttat ttatttattt    64260
atttatttat ttatttattt tgagacgcag tctcgctctg ttgcccaggt tggagggtgg    64320
tggcatgatc tcggctcact gcaagctccg cctcccgggt tcacgccatt ctcctgcctc    64380
actcttcctg agtagccgga ttacaggcgc acgcaccatg cctggctaat tattttgttt    64440
ttttagtaga gacagggttt cgccacgttg cccaggctgg tcttgaatcc ctggcctcaa    64500
gcgatccgcc cgcctcagcc tcccaaagtg ctgggattac aggcgtgagc caccgtgccc    64560
gcccagccta ggggtacatg aaactttttt tttttttttt ttgagacaga gtttcactct    64620
gtcctcaggc tggagtgcag tggcgtgatc tcggcgtact gcaatctccg cctcccggtt    64680
caagcgattc tcctgcctca gcctcccgag tagctgggat tgcaggcacg cgccaccaca    64740
cccagctaat ttttgtattt ttagtagaga cgggctttca ccatgtggga caggatggtc    64800
tcgatctcct gacctcgtga tccgcccgcc tcagcctccg aaagtgctgg gattacaggc    64860
ctgagccacc gtgcccagcc atgatgtttt gatacaggca tataacgtat aataatcaca    64920
tcagggtaaa tgatgtaacc atcacatcaa gcatttatcc tttgtgttac aaaaaaaaat    64980
ctaattatac tttcctactt attctttttt tttttttttt ttgagacgga gtctccctca    65040
gtcgcccagg ctggagtgca gtggcatgat ctcagttcac tgcaagctct gcctcctagc    65100
tctgcctcct gggttcatgc cattctcctg tctcagcctc gcgagtagct gggactacag    65160
gcgcctgcca ccgtgcccgg ctaatttttt tttttgtatt tttggtagag acagggtttc    65220
accgtgttag ccaggatggt ctcgatctcc tgacctcata atccgcccgt ctcggcctcc    65280
caaagtgctg ggattacagg catgagccac cgcccccagc ctatttattc ttaaatgtac    65340
aataaattat tgttgactcc agtcaccctg ctgtgctacc aaatacggat cttcttcatt    65400
ctatctaact gtatttctgt acctgttaac catctctcct ccacctcacc ccccaaaccc    65460
actacccttc tcagcctctg gtaaccatcc ttctactctc tatctctatg agttcaattg    65520
tattaatttt tagctccccg gccgggcacg gtggctcacg cctgtaatcc cagcacttca    65580
ggaggctgag gcaggtggat cacgaggtca ggagtttgag accagcctgg ccaacatggt    65640
ggaaccccat ctctactaaa aacacaaaaa ttagctgggc gtggtggtgg gcgcttgtag    65700
tcccagctac ttgggaggct gaggcaggag aatcgcttga aactgggagg cagaggttgc    65760
agtgagccaa gattgcgcca ctgcactcca gtctgggtga cagagtaaga ttccatcccg    65820
aaaaaaaaaa agtttagctc ccacaaataa gtgagaacac gtgaagtttc tctttctgtg    65880
cctcgcttgt ttcacttaac ataatgacct ccagttccat ccacgttgtt gctttgttat    65940
aaatgacagg atcttggtca ggcgcagtgg ctcatgcctg taatcccagc actttgggag    66000
gctgaggtgg actgatcatg aggtcaagag atcgagacca tcctggctaa cacagtgaaa    66060
ccccgtctct actaaaaata caagaaatta gccgggcgtg gtggtgggca cccatttccg    66120
ccccttctcg ggacgctgat gcacgacata ttacccatcc ccggaagact aatcctcccc    66180
cactctatat tgtacctctt cctttctcct ccacgcgatt ccccgagtaa cccgtcttcc    66240
ctccctcctc ggattacgct cacctttccg cttcaatcac gttgctccgt ccccttcccc    66300
```

```
attcgtacca ctcctcactt tcgtcttcct acccccacta tccctttttcg tcctctctat    66360
tccttactta ctcctccccc ttctcttcat acttcattcc ctccgctctt cccactcgcg    66420
ctcccacttt cacctagttg ccctcaccta cgttgccatc tcgccccttc ttcagctctc    66480
ggcctctcac ccatctgtcc tctctcttac ctctctcctc atctcgctca gacatctctc    66540
tagactatcc ctcactttac cttctcagtc gtcttcttcc tatccttcgt tctccatgat    66600
cttcacgtcg ccatctcttt tcgccccttt catatgtctc tcttcatgtt ctcactatca    66660
ttctcatgat cactatcgtt ctcactactt atcactcccc tctttcttca tcaattcctc    66720
tccgtcattc tcgtctctct cttacaaccg ccttccttgt gctatctaac tcaaccatgc    66780
ctctcctact ctctctctat cgcccctcca tcgcttatgc atcctcttct attgcacacc    66840
cgcccctcca tcgcttatgc atcctcttct attgcacacc gcccctccat cgcttatgca    66900
tcctcttcta ttgcacatcc tcttctattg cac                                66933
```

<210> 12
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 12
ctgagcggaa ttcgtgagac c 21

<210> 13
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 13
ttggtctcac gtattccgct cga 23

<210> 14
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 14
ctcgagaatt ctggatcctc 20

<210> 15
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 15
ttgaggatcc agaattctcg ag 22

<210> 16
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 16
tgtatgcgaa ttcgctgcgc g 21

<210> 17
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 17
ttcgcgcagc gaattcgcat aca 23

<210> 18
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 18
gtccactgaa ttctcagtga g 21

<210> 19
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 19
ttgtcactga gaattcagtg gac 23

<210> 20
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 20
gaatccgaat tcctggtcag c 21

<210> 21
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 21
ttgctgacca ggaattcgga ttc 23

<210> 22
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 22
cuacuacuac uactgagcgg aattcgtgag acc 33

<210> 23
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 23
cuacuacuac uactcgagaa ttctggatcc tc 32

<210> 24
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 24
cuacuacuac uatgtatgcg aattcgctgc gcg 33

<210> 25

<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 25
cuacuacuac uagtccactg aattctcagt gag 33

<210> 26
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 26
cuacuacuac uagaatccga attcctggtc agc 33

<210> 27
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 27
aactggaaga attcgcggcc gcaggaattt tttttttttt ttttt 45

<210> 28
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 28
aattcggcac gag 13

<210> 29
<211> 9
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 29
ctcgtgccg 9

<210> 30
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 30
gtacgacggc cagt 14

<210> 31
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 31
aacagctatg accatg 16

<210> 32
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 32
ccaagttctg agaagtcc 18

<210> 33
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 33
aatacctgaa accatacctg 20

<210> 34
<211> 57
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 34
agctgctcgt agctgtctct ccctggatca cgggtacatg tactggacag actgggt 57

<210> 35
<211> 56
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 35
tgagacgccc ggattgagcg ggcagggata gcttattccc tgtgccgcat tacggc 56

<210> 36
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 36
agctgctcgt agctgtctct ccctgga 27

<210> 37
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 37
gccgtaatgc ggcacaggga ataagct 27

<210> 38
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 38
gagaggctat atccctgggc 20

<210> 39
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 39
acagcacgtg tttaaagggg 20

<210> 40
<211> 163
<212> DNA
<213> Homo sapiens

<400> 40
actaaagcgc cgccgccgcg ccatggagcc cgagtgagct cggcgcgggc ccgtccggcc 60
gccggacaac atggaggcag ctccgcccgg gccgccgtgg ccgctgctgc tgctgctgct 120
gctgctgctg gcgctgtgcg gctgcccggc ccccgccgcg gcc 163

<210> 41
<211> 419
<212> DNA
<213> Homo sapiens

<400> 41
gccccacagc ctcgccgctc ctgctatttg ccaaccgccg ggacgtacgg ctggtggacg 60
ccggcggagt caagctggag tccaccatcg tggtcagcgg cctggaggat gcggccgcag 120
tggacttcca gttttccaag ggagccgtgt actggacaga cgtgagcgag gaggccatca 180
agcagaccta cctgaaccag acgggggccg ccgtgcagaa cgtggtcatc tccggcctgg 240
tctctcccga cggcctcgcc tgcgactggg tgggcaagaa gctgtactgg acggactcag 300
agaccaaccg catcgaggtg gccaacctca atggcacatc ccggaaggtg ctcttctggc 360
aggaccttga ccagccgagg gccatcgcct tggaccccgc tcacgggtaa accctgctg 419

<210> 42
<211> 221
<212> DNA
<213> Homo sapiens

<400> 42
ccccgtcaca ggtacatgta ctggacagac tggggtgaga cgccccggat tgagcgggca 60
gggatggatg gcagcacccg gaagatcatt gtggactcgg acatttactg gcccaatgga 120
ctgaccatcg acctggagga gcagaagctc tactgggctg acgccaagct cagcttcatc 180
caccgtgcca acctggacgg ctcgttccgg taggtaccca c 221

<210> 43
<211> 221
<212> DNA
<213> Homo sapiens

<400> 43
tccctgactg caggcagaag gtggtggagg gcagcctgac gcacccttc gccctgacgc 60
tctccgggga cactctgtac tggacagact ggcagacccg ctccatccat gcctgcaaca 120
agcgcactgg ggggaagagg aaggagatcc tgagtgccct atactcaccc atggacatcc 180
aggtgctgag ccaggagcgg cagccttttt gtgagtgccg g 221

<210> 44
<211> 156
<212> DNA
<213> Homo sapiens

<400> 44
tttctcagtc cacactcgct gtgaggagga caatggcggc tggtcccacc tgtgcctgct 60
gtccccaagc gagccttttt acacatgcgc ctgccccacg ggtgtgcaga tgcaggacaa 120
cggcaggacg tgtaaggcag gtgaggcggt gggacg 156

<210> 45
<211> 416
<212> DNA
<213> Homo sapiens

<400> 45
ctccacagga gccgaggagg tgctgctgct ggcccggcgg acggacctac ggaggatctc 60
gctggacacg ccggacttca ccgacatcgt gctgcaggtg gacgacatcc ggcacgccat 120
tgccatcgac tacgacccgc tagagggcta tgtctactgg acagatgacg aggtgcgggc 180
catccgcagg gcgtacctgg acgggtctgg ggcgcagacg ctggtcaaca ccgagatcaa 240
cgaccccgat ggcatcgcgg tcgactgggt ggcccgaaac ctctactgga ccgacacggg 300
cacggaccgc atcgaggtga cgcgcctcaa cggcacctcc cgcaagatcc tggtgtcgga 360
ggacctggac gagccccgag ccatcgcact gcaccccgtg atggggtaag acgggc 416

<210> 46
<211> 198
<212> DNA
<213> Homo sapiens

<400> 46
ttcttctcca gcctcatgta ctggacagac tggggagaga accctaaaat cgagtgtgcc 60
aacttggatg ggcaggagcg gcgtgtgctg gtcaatgcct ccctcgggtg gcccaacggc 120
ctggccctgg acctgcagga ggggaagctc tactggggag acgccaagac agacaagatc 180
gaggtgaggc tcctgtgg 198

<210> 47
<211> 244
<212> DNA
<213> Homo sapiens

<400> 47
ccgtcctgca ggtgatcaat gttgatggga cgaagaggcg gaccctcctg gaggacaagc 60
tcccgcacat tttcgggttc acgctgctgg gggacttcat ctactggact gactggcagc 120
gccgcagcat cgagcgggtg cacaaggtca aggccagccg ggacgtcatc attgaccagc 180
tgcccgacct gatggggctc aaagctgtga atgtggccaa ggtcgtcggt gagtccgggg 240
ggtc 244

<210> 48
<211> 313
<212> DNA
<213> Homo sapiens

<400> 48
gttcgcttcc aggaaccaac ccgtgtgcgg acaggaacgg ggggtgcagc cacctgtgct 60
tctgcacacc ccacgcaacc cggtgtggct gccccatcgg cctggagctg ctgagtgaca 120
tgaagacctg catcgtgcct gaggcctttt tggtcttcac cagcagagcc gccatccaca 180
ggatctccct cgagaccaat aacaacgacg tggccatccc gctcacgggc gtcaaggagg 240
cctcagccct ggactttgat gtgtccaaca accacatcta ctggacagac gtcagcctga 300
aggtagcgtg ggc 313

<210> 49
<211> 255
<212> DNA
<213> Homo sapiens

<400> 49
cctgctgcca gaccatcagc cgcgccttca tgaacgggag ctcggtggag cacgtggtgg 60
agtttggcct tgactacccc gagggcatgg ccgttgactg gatgggcaag aacctctact 120
gggccgacac tgggaccaac agaatcgaag tggcgcggct ggacgggcag ttccggcaag 180
tcctcgtgtg gagggacttg gacaacccga ggtcgctggc cctggatccc accaaggggt 240
aagtgtttgc ctgtc 255

<210> 50
<211> 210
<212> DNA
<213> Homo sapiens

<400> 50
gtgccttcca gctacatcta ctggaccgag tggggcggca agccgaggat cgtgcgggcc 60
ttcatggacg ggaccaactg catgacgctg gtggacaagg tgggccgggc caacgacctc 120
accattgact acgctgacca gcgcctctac tggaccgacc tggacaccaa catgatcgag 180
tcgtccaaca tgctgggtga gggccgggct 210

<210> 51
<211> 352
<212> DNA
<213> Homo sapiens

<400> 51
gtgttcatgc aggtcaggag cgggtcgtga ttgccgacga tctcccgcac ccgttcggtc 60
tgacgcagta cagcgattat atctactgga cagactggaa tctgcacagc attgagcggg 120
ccgacaagac tagcggccgg aaccgcaccc tcatccaggg ccacctggac ttcgtgatgg 180
acatcctggt gttccactcc tcccgccagg atggcctcaa tgactgtatg cacaacaacg 240
ggcagtgtgg gcagctgtgc cttgccatcc ccggcggcca ccgctgcggc tgcgcctcac 300
actacaccct ggaccccagc agccgcaact gcagccgtaa gtgcctcatg gt 352

<210> 52
<211> 225
<212> DNA
<213> Homo sapiens

<400> 52
gcctcctcta cgcccaccac cttcttgctg ttcagccaga aatctgccat cagtcggatg 60
atcccggacg accagcacag cccggatctc atcctgcccc tgcatggact gaggaacgtc 120
aaagccatcg actatgaccc actggacaag ttcatctact gggtggatgg gcgccagaac 180

atcaagcgag ccaaggacga cgggacccag gcaggtgccc tgtgg 225

<210> 53
<211> 235
<212> DNA
<213> Homo sapiens

<400> 53
ctttgtctta cagccctttg ttttgacctc tctgagccaa ggccaaaacc cagacaggca 60
gccccacgac ctcagcatcg acatctacag ccggacactg ttctggacgt gcgaggccac 120
caataccatc aacgtccaca ggctgagcgg ggaagccatg ggggtggtgc tgcgtgggga 180
ccgcgacaag cccagggcca tcgtcgtcaa cgcggagcga gggtaggagg ccaac 235

<210> 54
<211> 218
<212> DNA
<213> Homo sapiens

<400> 54
ccaccctccc gcaggtacct gtacttcacc aacatgcagg accgggcagc caagatcgaa 60
cgcgcagccc tggacggcac cgagcgcgag gtcctcttca ccaccggcct catccgccct 120
gtggccctgg tggtggacaa cacactgggc aagctgttct gggtggacgc ggacctgaag 180
cgcattgaga gctgtgacct gtcaggtacg cgccccgg 218

<210> 55
<211> 234
<212> DNA
<213> Homo sapiens

<400> 55
ggctgcttgc aggggccaac cgcctgaccc tggaggacgc caacatcgtg cagcctctgg 60
gcctgaccat ccttggcaag catctctact ggatcgaccg ccagcagcag atgatcgagc 120
gtgtggagaa gaccaccggg gacaagcgga ctcgcatcca gggccgtgtc gcccacctca 180
ctggcatcca tgcagtggag gaagtcagcc tggaggagtt ctgtacgtgg gggc 234

<210> 56
<211> 157
<212> DNA
<213> Homo sapiens

<400> 56
ttgtctttgc agcagcccac ccatgtgccc gtgacaatgg tggctgctcc cacatctgta 60
ttgccaaggg tgatgggaca ccacggtgct catgcccagt ccacctcgtg ctcctgcaga 120
acctgctgac ctgtggaggt aggtgtgacc taggtgc 157

<210> 57
<211> 272
<212> DNA
<213> Homo sapiens

<400> 57
gttctcctct gtccctcccc cagagccgcc cacctgctcc ccggaccagt ttgcatgtgc 60

cacaggggag atcgactgta tccccggggc ctggcgctgt gacggctttc ccgagtgcga 120
tgaccagagc gacgaggagg gctgccccgt gtgctccgcc gcccagttcc cctgcgcgcg 180
gggtcagtgt gtggacctgc gcctgcgctg cgacggcgag gcagactgtc aggaccgctc 240
agacgaggtg gactgtgacg gtgaggccct cc 272

<210> 58
<211> 134
<212> DNA
<213> Homo sapiens

<400> 58
tctccttgca gccatctgcc tgcccaacca gttccggtgt gcgagcggcc agtgtgtcct 60
catcaaacag cagtgcgact ccttccccga ctgtatcgac ggctccgacg agctcatgtg 120
tggtgagcca gctt 134

<210> 59
<211> 274
<212> DNA
<213> Homo sapiens

<400> 59
gtttgtctct ggcagaaatc accaagccgc cctcagacga cagcccggcc cacagcagtg 60
ccatcgggcc cgtcattggc atcatcctct ctctcttcgt catgggtggt gtctattttg 120
tgtgccagcg cgtggtgtgc cagcgctatg cgggggccaa cgggcccttc ccgcacgagt 180
atgtcagcgg gaccccgcac gtgcccctca atttcatagc cccgggcggt tcccagcatg 240
gcccccttcac aggtaaggag cctgagatat ggaa 274

<210> 60
<211> 164
<212> DNA
<213> Homo sapiens

<400> 60
cttccctgcc aggcatcgca tgcggaaagt ccatgatgag ctccgtgagc ctgatggggg 60
gccggggcgg ggtgcccctc tacgaccgga accacgtcac aggggcctcg tccagcagct 120
cgtccagcac gaaggccacg ctgtacccgc cggtgagggg cggg 164

<210> 61
<211> 130
<212> DNA
<213> Homo sapiens

<400> 61
ttggctctcc tcagatcctg aacccgccgc cctccccggc cacggacccc tccctgtaca 60
acatggacat gttctactct tcaaacattc cggccactgc gagaccgtac aggtaggaca 120
tcccctgcag 130

<210> 62
<211> 496
<212> DNA
<213> Homo sapiens

<400> 62
tcaaacattc cggccactgc gagaccgtac aggccctaca tcattcgagg aatggcgccc 60
ccgacgacgc cctgcagcac cgacgtgtgt gacagcgact acagcgccag ccgctggaag 120
gccagcaagt actacctgga tttgaactcg gactcagacc cctatccacc cccacccacg 180
ccccacagcc agtacctgtc ggcggaggac agctgcccgc cctcgcccgc caccgagagg 240
agctacttcc atctcttccc gccccctccg tcccccctgca cggactcatc ctgacctcgg 300
ccgggccact ctggcttctc tgtgcccctg taaatagttt taaatatgaa caaagaaaaa 360
aatatatttt atgatttaaa aaataaatat aattgggatt ttaaaaacat gagaaatgtg 420
aactgtgatg gggtgggcag ggctgggaga actttgtaca gtggagaaat atttataaac 480
ttaattttgt aaaaca 496

<210> 63
<211> 3081
<212> DNA
<213> Homo sapiens

<400> 63
cccgccagcc cagcccagcc caaccctact ccctccccac gccagggcag cagccgttgc 60
tcagagagaa ggtggaggaa gaaatccaga ccctagcacg cgcgcaccat catggaccat 120
tatgattctc agcaaaccaa cgattacatg cagccagaag aggactggga ccgggacctg 180
ctcctggacc cggcctggga gaagcagcag agaaagacat tcacggcatg gtgtaactcc 240
cacctccgga aggcggggac acagatcgag aacatcgaag aggacttccg ggatggcctg 300
aagctcatgc tgctgctgga ggtcatctca ggtgaacgct tggccaagcc agagcgaggc 360
aagatgagag tgcacaagat ctccaacgtc aacaaggccc tggatttcat agccagcaaa 420
ggcgtcaaac tggtgtccat cggagccgaa gaaatcgtgg atgggaatgt gaagatgacc 480
ctgggcatga tctggaccat catcctgcgc tttgccatcc aggacatctc cgtggaagag 540
acttcagcca aggaagggct gctcctgtgg tgtcagagaa agacagcccc ttacaaaaat 600
gtcaacatcc agaacttcca cataagctgg aaggatggcc tcggcttctg tgctttgatc 660
caccgacacc ggcccgagct gattgactac gggaagctgc ggaaggatga tccactcaca 720
aatctgaata cggcttttga cgtggcagag aagtacctgg acatccccaa gatgctggat 780
gccgaagaca tcgttggaac tgcccgaccg gatgagaaag ccatcatgac ttacgtgtct 840
agcttctacc acgccttctc tggagcccag aaggcggaga cagcagccaa tcgcatctgc 900
aaggtgttgg ccgtcaacca ggagaacgag cagcttatgg aagactacga gaagctggcc 960
agtgatctgt tggagtggat ccgccgcaca atcccgtggc tggagaaccg ggtgcccgag 1020
aacaccatgc atgccatgca acagaagctg gaggacttcc gggactaccg gcgcctgcac 1080
aagccgccca aggtgcagga gaagtgccag ctggagatca acttcaacac gctgcagacc 1140
aagctgcggc tcagcaaccg gcctgccttc atgccctctg agggcaggat ggtctcggac 1200
atcaacaatg cctggggctg cctggagcag gtggagaagg gctatgagga gtggttgctg 1260
aatgagatcc ggaggctgga gcgactggac cacctggcag agaagttccg gcagaaggcc 1320
tccatccacg aggcctggac tgacggcaaa gaggccatgc tgcgacagaa ggactatgag 1380
accgccaccc tctcggagat caaggccttg ctcaagaagc atgaggcctt cgagagtgac 1440
ctggctgccc accaggaccg tgtggagcag attgccgcca tcgcacagga gctcaatgag 1500
ctggactatt atgactcacc cagtgtcaac gcccgttgcc aaaagatctg tgaccagtgg 1560
gacaatctgg gggccctaac tcagaagcga agggaagctc tggagcggac cgagaaactg 1620
ctggagacca ttgaccagct gtacttggag tatgccaagc gggctgcacc cttcaacaac 1680
tggatggagg gggccatgga ggacctgcag gacaccttca ttgtgcacac cattgaggag 1740
atccagggac tgaccacagc ccatgagcag ttcaaggcca ccctccctga tgccgacaag 1800
gagcgcctgg ccatcctggg catccacaat gaggtgtcca agattgtcca gacctaccac 1860
gtcaatatgg cgggcaccaa cccctacaca accatcacgc ctcaggagat caatggcaaa 1920
tgggaccacg tgcggcagct ggtgcctcgg agggaccaag ctctgacgga ggagcatgcc 1980
cgacagcagc acaatgagag gctacgcaag cagtttggag cccaggccaa tgtcatcggg 2040
ccctggatcc agaccaagat ggaggagatc gggaggatct ccattgagat gcatgggacc 2100
ctggaggacc agctcagcca cctgcggcag tatgagaaga gcatcgtcaa ctacaagcca 2160
aagattgatc agctggaggg cgaccaccag ctcatccagg aggcgctcat cttcgacaac 2220
aagcacacca actacaccat ggagcacatc cgtgtgggct gggagcagct gctcaccacc 2280
atcgccagga ccatcaatga ggtagagaac cagatcctga cccgggatgc caagggcatc 2340

```
agccaggagc agatgaatga gttccgggcc tccttcaacc actttgaccg ggatcactcc    2400
ggcacactgg gtcccgagga gttcaaagcc tgcctcatca gcttgggtta tgatattggc    2460
aacgaccccc agggagaagc agaatttgcc cgcatcatga gcattgtgga ccccaaccgc    2520
ctgggggtag tgacattcca ggccttcatt gacttcatgt cccgcgagac agccgacaca    2580
gatacagcag accaagtcat ggcttccttc aagatcctgg ctggggacaa gaactacatt    2640
accatggacg agctgcgccg cgagctgcca cccgaccagg ctgagtactg catcgcgcgg    2700
atggccccct acaccggccc cgactccgtg ccaggtgctc tggactacat gtccttctcc    2760
acggcgctgt acggcgagag tgacctctaa tccaccccgc ccggccgccc tcgtcttgtg    2820
cgccgtgccc acagatgtga aatgaatgta atctaataga agcctaatca gcccaccatg    2880
ttctccactg aaaaatcctc tttctttggg gtttttcttt ctttcttttt tgattttgca    2940
ctggacggtg acgtcagcct gtacaggctc ccaggggtgg cgtcaaatgc tattgaaatt    3000
gcgctgaatc gtatgctttt tccttttgat aaataaacaa tgtaaaaatg tttcaaaaac    3060
ctaataaaat aaataaatac g                                              3081


<210> 64
<211> 1324
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(1324)
<223> n = A,T,C or G

<400> 64
ggccgcccgg cgcccccagc agnccgagcc ggggcgcaca gncggggcgc agcccgcgcc      60
ccccgccgcg attgacatga tgtttccaca aagcaggcat tcgggctcct cgcacctacc     120
ccagcaactc aaattcacca cctcggactc ctgcgaccgc atcaaagacg aatttcagct     180
actgcaagct cagtaccaca gcctcaagct cgaatgtgac aagttggcca gtgagaagtc     240
agagatgcag cgtcactatg tgatgtacta cgagatgtcc tacggcttga acatcgagat     300
gcacaaacag gctgagatcg tcaaaaggct gaacgggatt tgtgcccagg tcctgcccta     360
cctctcccaa gagcaccagc agcaggtctt gggagccatt gagagggcca agcaggtcac     420
cgctcccgag ctgaactcta tcatccgaca gcagctccaa gcccaccagc tgtcccagct     480
gcaggccctg gccctgccct tgaccccact acccgtgggg ctgcagccgc cttcgctgcc     540
ggcggtcagc gcaggcaccg gcctcctctc gctgtccgcg ctgggttccc aggcccacct     600
ctccaaggaa gacaagaacg ggcacgatgg tgacacccac caggaggatg atggcgagaa     660
gtcggattag caggggggccg ggacagggag gttgggaggg gggacagagg ggagacagag     720
gcacggagag aaaggaatgt ttagcacaag acacagcgga gctcgggatt ggctaatctc     780
ccatagtatt tatggtggcg ccggcggggc cccagcccag cttgcaggcc acctctagct     840
ttcttcctac cccattccgg cttccctcct cctcccctgc agcctggtta ggtggatacc     900
tgccctgaca tgtgaggcaa gctaaggcct ggagggtcag atgggagacc aggtcccaag     960
ggagcaagac ctgcgaagcg cagcagcccc ggcccttccc ccgttttgaa catgtgtaac    1020
cgacagtctg ccctgggcca cagccctctc accctggtac tgcatgcacg caatgctagc    1080
tgcccctttc ccgtcctggg caccccgagt ctcccccgac cccgggtccc aggtatgctc    1140
ccacctccac ctgccccact caccacctct gctagttcca gacacctcca cgcccacctg    1200
gtcctctccc atcgcccaca aaaggggggg cacgagggac gagcttagct gagctgggag    1260
gagcagggtg agggtgggcg acccaggatt ccccctcccc ttcccaaata aagatgaggg    1320
tact                                                                 1324

<210> 65
<211> 2377
<212> DNA
<213> Homo sapiens

<400> 65
ggtgacaaag agccaacaga gacaatagga gacttgtcaa tttgtcttga tgggctacag      60
```

ttagagtctg aagttgttac caatggtgaa actacatgtt cagaaagtgc ttctcagaat 120
gatgatggct ccagatccaa ggatgaaaca agagtgagca caaatggatc agatgaccct 180
gaagatgcag gagctggtga aaataggaga gtcagtggga ataattctcc atcactctca 240
aatggtggtt ttaaaccttc tagacctcca agaccttcac gaccaccacc acccaccccca 300
cgtagaccag catctgtcaa tggttcacca tctgccactt ctgaaagtga tgggtctagt 360
acaggctctc tgccgccgac aaatacaaat acaaatacat ctgaaggagc aacatctgga 420
ttaataattc ctcttactat atctggaggc tcaggcccta ggccattaaa tcctgtaact 480
caagctccct tgccacctgg ttgggagcag agagtggacc agcacgggcg agtttactat 540
gtagatcatg ttgagaaaag aacaacatgg gatagaccag aacctctacc tcctggctgg 600
gaacggcggg ttgacaacat gggacgtatt tattatgttg accatttcac aagaacaaca 660
acgtggcaga ggccaacact ggaatccgtc cggaactatg aacaatggca gctacagcgt 720
agtcagcttc aaggagcaat gcagcagttt aaccagagat tcatttatgg gaatcaagat 780
ttatttgcta catcacaaag taaagaattt gatcctcttg gtccattgcc acctggatgg 840
gagaagagaa cagacagcaa tggcagagta tatttcgtca accacaacac acgaattaca 900
caatgggaag accccagaag tcaaggtcaa ttaaatgaaa agcccttacc tgaaggttgg 960
gaaatgagat tcacagtgga tggaattcca tattttgtgg accacaatag aagaactacc 1020
acctatatag atccccgcac aggaaaatct gccctagaca atggacctca gatagcctat 1080
gttcgggact tcaaagcaaa ggttcagtat ttccggttct ggtgtcagca actggccatg 1140
ccacagcaca taaagattac agtgacaaga aaaacattgt ttgaggattc ctttcaacag 1200
ataatgagct tcagtcccca agatctgcga agacgtttgt gggtgatttt tccaggagaa 1260
gaaggtttag attatggagg tgtagcaaga gaatggttct ttcttttgtc acatgaagtg 1320
ttgaacccaa tgtattgcct gtttgaatat gcagggaagg ataactactg cttgcagata 1380
aaccccgctt cttacatcaa tccagatcac ctgaaatatt ttcgttttat tggcagattt 1440
attgccatgg ctctgttcca tgggaaattc atagacacgg gtttttcttt accattctat 1500
aagcgtatct tgaacaaacc agttggactc aaggatttag aatctattga tccagaattt 1560
tacaattctc tcatctgggt taaggaaaac aatattgagg aatgtgattt ggaaatgtac 1620
ttctccgttg acaaagaaat tctaggtgaa attaagagtc atgatctgaa acctaatggt 1680
ggcaatattc ttgtaacaga agaaaataaa gaggaataca tcagaatggt agctgagtgg 1740
aggttgtctc gaggtgttga agaacagaca caagcttttct ttgaaggctt taatgaaatt 1800
cttccccagc aatatttgca atactttgat gcaaaggaat tagaggtcct tttatgtgga 1860
atgcaagaga ttgatttgaa tgactggcaa agacatgcca tctaccgtca ttatgcaagg 1920
accagcaaac aaatcatgtg gttttggcag tttgttaaag aaattgataa tgagaagaga 1980
atgagacttc tgcagtttgt tactggaacc tgccgattgc cagtaggagg atttgctgat 2040
ctcatgggga gcaatggacc acagaaattc tgcattgaaa aagttgggaa agaaaattgg 2100
ctacccagaa gtcatacctg ttttaatcgc ctggacctgc caccatacaa gagctatgag 2160
caactgaagg aaaagctgtt gtttgccata gaagaaacag aaggatttgg acaagagtaa 2220
cttctgagaa cttgcaccat gaatgggcaa gaacttattt gcaatgtttg tccttctctg 2280
cctgttgcac atcttgtaaa attggacaat ggctctttag agagttatct gagtgtaagt 2340
aaattaatgt tctcatttaa aaaaaaaaaa aaaaaaa 2377

<210> 66
<211> 1295
<212> DNA
<213> Homo sapiens

<400> 66
gggccgccgc ccacccgggc cttgcctcta cctcagtcgt tgccccccga ttttcggctg 60
gagcccacgg ccccggccct cagccccccgc tctagcttcg ccagtagctc ggccagcgac 120
gcgagcaagc cgtccagccc ccggggcagc ctgctgctgg acggggcggg ggctggcgga 180
gctggaggta gccggccctg cagcaatcgc accagcggca tcagcatggg ctacgaccag 240
cgccacggga gcccccttgcc agcgggggccg tgcctgtttg gcccacccct ggccggagca 300
ccggcaggct attctcccgg aggggtcccg tccgcctacc cggagctcca cgccgccctg 360
gaccgattgt acgctcagcg gcccgcgggg ttcggctgcc aggaaagccg ccactcgtat 420
cccccggccc tgggcagccc tggagctcta gccgggggccc gagtgggagc ggcgggggccc 480
ttggagagac gggggggcgca acccggacga cactctgtga ccggctacgg ggactgcgcc 540
gtgggcgccc ggtaccagga cgagctaaca gctttgcttc gcctgacggt gggcaccggt 600
gggcgagaag ccggagcccg cggagaaccc tcggggattg agccgtcggg tctggaggag 660

```
ccaccaggtc ctttcgttcc ggaggccgcc cgggcccgga tgcgggagcc agaggccagg      720
gaggactact tcggcacctg tatcaagtgc aacaaaggca tctatgggca gagcaatgcc      780
tgccaggccc tggacagcct ctaccacacc cagtgctttg tttgctgctc ttgtgggcga      840
actttgcgtt gcaaggcttt ctacagtgtc aatggctctg tgtactgtga ggaagattat      900
ctgttttcag ggtttcagga ggcagctgag aaatgctgtg tctgtggtca cttgattttg      960
gagaagatcc tacaagcaat ggggaagtcc tatcatccag gctgtttccg atgcattgtt     1020
tgcaacaagt gcctggatgg catccccttc acagtggact tctccaacca agtatactgt     1080
gtcaccgact accacaaaaa ttatgctcct aagtgtgcag cctgtggcca acccatcctc     1140
ccctctgagg gctgtgagga catcgtgagg gtgatatcca tggaccggga ttatcacttt     1200
gagtgctacc actgtgagga ctgccggatg cagctgagtg atgaggaagg ctgctgctgt     1260
ttccctctgg atgggcactt gctctgccat ggttg                                1295
```

<210> 67
<211> 3411
<212> DNA
<213> Homo sapiens

<400> 67

```
gggcccgggg tcccgccacc accgcgcgcg ggacagattg attcactttg gagctgtaag       60
tactgatgta ttagggtgca gcgctcattg ttcattgacg cagagtccca aaatgaatat      120
ccaagagcag ggtttcccct tggacctcgg agcaagtttc accgaagatg ctcccccgacc     180
cccagtgcct ggtgaggagg gagaactggt gtccacagac ccgaggcccg ccagctacag      240
tttctgctcc gggaaaggtg ttggcattaa aggtgagact tcgacggcca ctccgaggcg      300
ctcggatctg gacctggggt atgagcctga gggcagtgcc tcccccaccc caccatactt      360
gaagtgggct gagtcactgc attccctgct ggatgaccaa gatgggataa gcctgttcag      420
gactttcctg aagcaggagg gctgtgccga cttgctggac ttctggtttg cctgcactgg      480
cttcaggaag ctggagccct gtgactcgaa cgaggagaag aggctgaagc tggcgagagc      540
catctaccga aagtacattc ttgataacaa tggcatcgtg tcccggcaga ccaagccagc      600
caccaagagc ttcataaagg gctgcatcat gaagcagctg atcgatcctg ccatgtttga      660
ccaggcccag accgaaatcc aggccactat ggaggaaaac acctatccct ccttccttaa      720
gtctgatatt tatttggaat atacgaggac aggctcggag agccccaaag tctgtagtga      780
ccagagctct gggtcaggga cagggaaggg catatctgga tacctgccga ccttaaatga      840
agatgaggaa tggaagtgtg accaggacat ggacgaggac gatggcagag acgctgctcc      900
ccccggaaga ctccctcaga agctgctcct ggagacagct gccccgaggg tctcctccag      960
tagacggtac agcgaaggca gagagttcag gtatggatcc tggcgggagc cagtcaaccc     1020
ctattatgtc aatgccggct atgccctggc cccagccacc agtgccaacg acagcgagca     1080
gcagagcctg tccagcgatg cagacaccct gtccctcacg gacagcagcg tggatgggat     1140
ccccccatac aggatccgta agcagcaccg cagggagatg caggagagcg cgcaggtcaa     1200
tgggcgggtg cccctacctc acattccccg cacgtaccgg gtgccgaagg aggtccgcgt     1260
ggagcctcag aagttcgcgg aggagctcat ccaccgcctg gaggctgtgc agcgcacgcg     1320
ggaggccgag gagaagctgg aggagcggct gaagcgcgtg cgcatggagg aggaaggtga     1380
ggacggcgat ccatcgtcag ggcccccagg gccgtgtcac aagctgcctc ccgcccccgc     1440
ttggcaccac ttcccgcccc gcttgtgttg gacatgggct tgtgccgggc tccgggatgc     1500
acacgaggag aaccctgaga gcatcctgga cgagcacgta cagcgtgtgc tgaggacaac     1560
tggccgccag tcgcctgggc ctggccatcg ctccccggac agtgggcacg tggccaagat     1620
gccagtggca ctgggggggtg ccgcctcggg gcacgggaag cacgtaccca agtcaggggc    1680
gaagctggac gcggccggcc tgcaccacca ccgacacgtc caccaccacg tccaccacag     1740
cacagcccgg cccaaggagc aggtggaggc cgaggccacc cgcagggccc agagcagctt     1800
cgcctggggc ctggaaccac acagccatgg ggcaaggtcc cgaggctact cagagagtgt     1860
tggcgctgcc cccaacgcca gtgatggcct cgcccacagt gggaaggtgg gcgttgcgtg     1920
caaaagaaat gccaagaagg ctgagtcggg gaagagcgcc agcaccgagg tgccaggtgc     1980
ctcggaggat gcggagaaga accagaaaat catgcagtgg atcattgagg gggaaaagga     2040
gatcagcagg caccgcagga ccggccacgg gtcttcgggg acgaggaagc cacagcccca     2100
tgagaactcc agaccyttgt cccttgagca cccctgggcc ggccctcagc tccggacctc     2160
cgtgcagccc tcccacctct tcatccaaga ccccaccatg ccaccccacc cagctcccaa     2220
ccccctaacc cagctggagg aggcgcgccg acgtctggag gaggaagaaa agagagccag     2280
ccgagcaccc tccaagcaga ggtatgtgca ggaggttatg cggcggggac gcgcctgcgt     2340
```

caggccagcg tgcgcgccgg tgctgcacgt ggtaccagcc gtgtcggaca tggagctctc 2400
cgagacagag acaagatcgc agaggaaggt gggcggcggg agtgcccagc cgtgtgacag 2460
catcgttgtg gcgtactact tctgcgggga acccatcccc taccgcaccc tggtgagggg 2520
ccgcgctgtc accctgggcc agttcaagga gctgctgacc aaaaagggca gctacagata 2580
ctacttcaag aaagtgagcg acgagtttga ctgtggggtg gtgtttgagg aggttcgaga 2640
ggacgaggcc gtcctgcccg tctttgagga gaagatcatc ggcaaagtgg agaaggtgga 2700
ctgataggct ggtgggctgg ccgctgtgcc aggcgaggcc cttggcgggc acgggtgtca 2760
cggccaggca gatgacctcg tactcaggag cccgatgggg aacagtgttg ggtgtaccac 2820
ccatccctgt ggtctacccg tgtctagagg caggtagggg gtccctccaa gtggtccaca 2880
agcttctgtc ctgcccccaa ggaggcagcc tggaccactc ctcatagcaa tacttggagg 2940
gcccagccca agtgaggcag ccgaggtccc tgctgccagc ttcaggtgac cccccccat 3000
cccccggcac ctcccttggg cacgtgtgct gggatctact ttccctctgg gatttgccca 3060
cgtacccagg tctggctggg gcccaggccc ggatgcagag gcctgcaggg cctctgtcaa 3120
ttgtacgcgc caccaagtgc cttcaacaca gcttgtctct tgcctgccac tgtgtgaatc 3180
ggcgacggag cactgcacct gcctccagcc gccggctgtg cagtcctggg tcctcctttc 3240
tgagggcccg tgtaaatatg tacatttctc aggctagggc cagcaggggc tgcccgagtc 3300
tgtttttcat gcgatgacac.ttgtacaatt atcttttcaa aggtacttgg ataataatga 3360
aataaaactg tttttgaacc tgaataaaaa aaaaaaaaaa aaaaaaaaaa a 3411

<210> 68
<211> 3140
<212> DNA
<213> Homo sapiens

<400> 68
ggctgcgagt acctccatgg tcccggtggc tgtgacggcg gcagtggcgc ctgtcctgtc 60
cataaacagc gatttctcag atttgcggga aattaaaaag caactgctgc ttattgcggg 120
ccttacccgg gagcggggcc tactacacag tagcaaatgg tcggcggagt tggctttctc 180
tctccctgca ttgcctctgg ccgagctgca accgcctccg cctattacag aggaagatgc 240
ccaggatatg gatgcctata ccctggccaa ggcctacttt gacgttaaag agtatgatcg 300
ggcagcacat ttcctgcatg gctgcaatag caagaaagcc tattttctgt atatgtattc 360
cagatatctg tctggagaaa aaaagaagga cgatgaaaca gttgatagct taggcccccct 420
ggaaaaagga caagtgaaaa atgaggcgct tagagaattg agagtggagc tcagcaaaaa 480
acaccaagct cgagaacttg atggatttgg actttatctg tatggtgtgg tgcttcgaaa 540
actggacttg gttaaagagg ccattgatgt gtttgtggaa gctactcatg ttttgccctt 600
gcattgggga gcctggttag aactctgtaa cctgatcaca gacaaagaga tgctgaagtt 660
cctgtctttg ccagacacct ggatgaaaga gtttttctg gctcatatat acacagagtt 720
gcagttgata gaggaggccc tgcaaaagta tcagaatctc attgatgtgg gcttctctaa 780
gagctcgtat attgtttccc aaattgcagt tgcctatcac aatatcagag atattgacaa 840
agccctctcc atttttaatg agctaaggaa acaagaccct tacaggattg aaaatatgga 900
cacattctcc aaccttcttt atgtcaggag catgaaatcg gagttgagtt atctggctca 960
taacctctgt gagattgata aataccgtgt agaaacgtgc tgtgtaattg gcaattatta 1020
cagtttacgt tctcagcatg agaaagcagc cttatatttc cagagagccc tgaaattaaa 1080
tcctcggtat cttggtgcct ggacactaat gggacatgag tacatggaga tgaagaacac 1140
gtctgctgct atccaggctt atagacatgc cattgaggtc aacaaacggg actacagagc 1200
ttggtatggc ctcgggcaga cctatgaaat ccttaagatg ccattttact gcctttatta 1260
ttatagacgg gcccaccagc ttcgacccaa tgattctcgc atgctggttg ctttaggaga 1320
atgttacgag aaactcaatc aactagtgga agccaaaaag tgttattgga gagcttacgc 1380
cgtgggagat gtggagaaaa tggctctggt gaaactggca aagcttcatg aacagttgac 1440
tgagtcagaa caggctgccc agtgttacat caaatatatc caagatatct attcctgtgg 1500
ggaaatagta gaacacttgg aggaaagcac tgcctttcgc tatctggccc agtactattt 1560
taagtgcaaa ctgtgggatg aagcttcaac ttgtgcacaa aagtgttgtg catttaatga 1620
tacccgggaa gaaggtaagg ccttactccg gcaaatccta cagcttcgga accaaggcga 1680
gactcctacc accgaggtgc ctgctccctt tttcctacct gcttcactct ctgctaacaa 1740
tacccccaca cgcagagttt ctccactcaa cttgtcttct gtcacgccat agttggctac 1800
tctcaagcca gcacattgtt agacccatct taattaagcc ttacctccat gtaaagaaca 1860
gcacgtctgt tccaaggacc tcagctcttc ttgtttctac agatggcaac agctccatag 1920

ggacagcttg tataattacc ttcagaggcc aactgacaga atcctggcag gaacagacat 1980
tatcttgcca gttagaagta cttctgtctc acttatgtcc aaagagtggc tatagatctt 2040
ggccttcttc cctgaatgct tttttttttt ggcccccaag aaagtcccctt ttatagcact 2100
ttagcacagg caatgctaca ggaacaaagt ttcaatgctg ctgagagtga aagaaaggag 2160
gaaagtctgc cactctaccc tgagctggca gtagggcact gagtaccctt aggaagaagt 2220
cagagcaatg gatacaaatg accttgctct tggatttgct gagcatgatc cctattctga 2280
tgtcagagat taggtttaaa tggaatagag ctatccattt gttcttactc tctagggaga 2340
caatcttcca aaacagtttt gggggggtct tctaaagctt tcaaattgga agtaacttta 2400
ttcaactaga gttgaataaa agaagggcaa aaataatctc acagagcttg gaactgctga 2460
tagcccttac tgagggcaaa agatggctat attgttagct atactcctac caaagcaagc 2520
aaggagatag gattatagat aatttcacgg acatttggaa ataacattgg tgattataca 2580
gacaagaata aactcacttc aagctggtct gttttaataa attttcaacg taattgtcta 2640
tttttttccc tcccatctgc aacagaatac attttttca gcctttatct agatgaggta 2700
aagggaatca ttcttatggt gctcttggag agtttcaggc ctgtgcatgt gtgtacagca 2760
ggaggtaata tgctataatg tctgctgtaa tatatttgca cagtagatgc tatggatcat 2820
tctgagctca gggtccagac tttattctta ttcccagaat tttgtgttac gtttttacct 2880
cctaacatat gacacttcat cttatattaa ggaaggttta gaatatctaa tacgacttga 2940
attcatttgt tactaagcct tctcaggcaa gctgtatact agttactggt ctccactgcc 3000
atgccttttc aaggttccca tggtccagaa tgatgtttga ttcttaattt ttctgtccct 3060
tttataattt gttttaatga ttttgctaca tttggaattc aataaaaaat gtgaacaata 3120
ataaaaaaaa aaaaaaaaaa 3140

<210> 69
<211> 3513
<212> DNA
<213> Homo sapiens

<400> 69
ccgtgtacca ggtgctgcta gtgggaagca cgctgctgaa ggaagtgcct tccgggctgc 60
agctggagca gttgccttct cagagcctgc tgacccacat cccaacggcg gggctgccca 120
cttcgctagg aggaggcctg ccttactgcc accaggcctg gctggatttc cgaaggcggc 180
tggaagctct actacagaac tgccaggcag cttgtgccct gctccagggg gccatcgaaa 240
gtgtgaaggc tgtgccccag cccatggagc ctggggaggt cggtcagctg ctacagcaga 300
cagaggtcct gatgcagcag gtgctagact cgccatggct ggcatggcta caatgccagg 360
ggggccggga gctgacatgg ctgaagcaag aggtcccaga ggtgaccctg agcccagact 420
acaggacggc aatggacaag gctgacgagc tatatgaccg ggtggatgga ttgctgcacc 480
aactgaccct gcagagcaac cagcgaatac aggccctaga gttggtccaa acactggagg 540
cccgggaaag cggactgcac cagattgaag tgtggctgca gcaggtgggc tggccagcac 600
tggaggaggc tggggagccc tcgctggaca tgctgctcca ggcccaaggc tcttttcagg 660
agctgtacca ggttgcccag gagcaggtca ggcaagggga gaagtttctg cagccgctga 720
ctggctggga ggcggctgaa ctggaccccc ctggggcacg ctttctggcc ctgcgagccc 780
agctgactga attctctagg gctttggccc agcggtgcca gcggctggcg gatgctgaga 840
ggctgtttca gctcttcagg gaggccttga cgtgggctga ggaggggcag cgagtgttgg 900
cagagctgga gcaggaacgc ccggggggttg tgttgcagca gctgcagctg cactggacca 960
ggcaccctga cttgcctcct gcccacttcc gaaagatgtg ggctctggcc acggggctgg 1020
gctcagaggc catccgccag gagtgccgct gggcctgggc gcggtgccag gacacctggc 1080
tggccctgga ccaaaagctt gaggcttcac tgaagctacc accggtgggc agcacagcta 1140
gcctgtgtgt cagccaggtc cccgctgcac ctgcccaccc tcccctgagg aaggcctaca 1200
gcttcgatcg gaatctgggg cagagtctca gtgaacctgc ctgccactgc caccatgcgg 1260
ccactattgc tgcctgccgc agaccagagg ctggaggagg tgccctgccc caggcatccc 1320
ctactgtgcc tccaccaggc agctctgacc ccaggagcct caacaggcta cagctggtgc 1380
tggcagagat ggtggccacg gagcgggagt atgtccgggc tctagagtac actatggaga 1440
actatttccc cgagctggat cgccccgatg tgccccaggg cctccgcggt cagcgtgccc 1500
acctctttgg caacctggag aagctgcggg acttccactg ccacttcttc ctgcgtgagc 1560
tggaggcttg cacccggcac ccaccacgag tggcctatgc cttcctgcgc catagggtgc 1620
agtttgggat gtacgcgctc tacagcaaga ataagcctcg ctccgatgcc ctgatgtcaa 1680
gctatgggca caccttcttc aaggacaagc agcaagcact gggggaccac ctggacctgg 1740

```
cctcctacct gctaaagccc atccagcgca tgggcaagta cgcactgctg ctgcaggagc    1800
tggcacgggc ctgcggggc cccacgcagg agctcagtgc gctgcgggag gcccagagcc    1860
ttgtgcactt ccagctgcgg cacggaaacg acctgctggc catggacgcc atccagggct    1920
gtgatgttaa cctcaaggaa caggggcagc tggtgcgaca ggatgagttt gtggtgcgca    1980
ctgggcgcca caagtccgtg cgccgcatct tcctttttga ggagctgctg ctcttcagca    2040
agcctcgcca tgggcccaca ggggttgaca catttgccta caagcgctcc ttcaagatgg    2100
cagaccttgg tctcactgag tgctgtggga acagcaacct gcgcttcgag atctggttcc    2160
gccgccgcaa ggccagggac acctttgtgc tgcaggcctc cagcctggct atcaagcagg    2220
cctggacagc tgacatctcc cacctgcttt ggaggcaggc cgtccacaac aaggaggtgc    2280
gcatggctga gatggtgtcc atgggtgtgg ggaacaaggc cttccgagac attgctccca    2340
gcgaggaagc catcaacgac cgcaccgtca actatgtcct gaagtgccga gaagttcgct    2400
ctcgggcgtc cattgccgta gccccgtttg accatgacag cctctacctg gggggcctcga   2460
actcccttcc tggagaccct gcctcttgct ctgttctggg gtccctcaac ctgcacctgt    2520
acagagaccc agctcttctg ggtctccgct gtcccctgta tcccagcttc ctagaggaag    2580
cagcactgga ggctgaggca gagctgggcg gccagccctc tttgactgct gaggactcag    2640
agatctcgtc ccaatgccca tcagccagtg gctccagtgg ctctgacagc agctgtgtgt    2700
cagggcaggc cctgggtagg ggcctggagg acttaccctg tgtctgagcc cgggactgga    2760
cgagcagtag atccagcagc ctgcagctcc aaggaacatt gcctctctgg atctgctgtg    2820
accagggtgt ggctgacacc tgggctacct ccaacctaca tgtgcaacgc tgttgactac    2880
cctttctgat gtgtgtggcc attggactaa ctggcacggg gcctctctag ggaagtctgg    2940
ttgtagagcc tgaataggct cctggcccca tgaccccttc tcctgtcccc agctcccatc    3000
ccagttgtgg gttaagaata ggctagagca gacattgggt gtttccatgc tgtaggctgg    3060
tgggggacca tgtgcctcta ggcagtgact agggtgcccc cacccctcag gaagaacaca    3120
ggtgggctcc tagcagctga tccccaatgc ctggccttaa agccgagctc agttaccata    3180
gggacaggtc cacctctact gggccctcat gcttgccttt cctggccccc aggcccagcc    3240
ccttttttact ggggcagttt cgttattttg acttgatgcc ttttgaataa ctttcaatag   3300
aattgtctaa aattatctta ctggttgtta ggcctttggt gtctcagaga aggagtctag    3360
gtctttgatg tgtgatttaa tcttttattt gtttataata aaaaatagac tgatttgtaa    3420
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa    3480
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa                                 3513
```

<210> 70
<211> 3597
<212> DNA
<213> Homo sapiens

<400> 70

```
catgccagtt acttcctcag gaaaatattt tcttgccttc ttctttcagt atggttttaa      60
atttgggaac agtggataac ccaagtgtcc cacaggccaa ggtatattcc aatggcagca     120
tgatccctgc acccaaagcc agcccctaaa gcctacccct tgtgcacccg cagcctggta     180
agtgagcttg gctgcttgtg aggagctaca agtgaaagag aagttatttt aaataaatcc     240
caaagtttga ggcagactgt ccaggactgt tcccaggaag aagcaggagt tacccacagg     300
aaaagtctct gacctggtcc cctcaggccc agctacctgc gcccaccagc agtgaaggtt     360
gatgtactgg cccagcatct ccacctcccc catgcaacca ggtccctggt accgtgtctc     420
ccgttgcatg tctggcttct gcctgtgctc ctcctgccac gagcatcctc cctgtccctc     480
ctcattccac cgtgtctctc ctgcacacat agcctctgtc ccagggcgat ttatccactt     540
gagtacagga gctgctcaga cctctcagcc cagccctctg tgactgcccc agccccatcc     600
taccccaccc aaagctgcct tcctggctgt aggagctccc tcgtctagcc aaggccctat     660
gggtccccat ccgaggatcc acaagcaatg acttcccaaa tgacctccac tgcaagaaga     720
atccttacca ctgtttccag agccgtgaac gatgctgtga tggcccaggt ctcagcacca     780
ccctctgtga cctaaaaaga aaagctcaat ttccatctgt cttctttccc aggaccaagg     840
ggacacagta atgtgaagtc aaatacttaa ccgagcaaag ggccagtgtt gttatcagtc     900
aaggacaaac ctcccacctc acagacagcc aagcagtgag ggaaagacag acagacatag     960
gtaggaaggt gctctgcagg cacaaggccc agagaagccc ctctccggga acttcccctg    1020
ctccttccag gaacagtgag cccagtgagc agtcccagcc agctcttcaa ggccttcaag    1080
gggtctttcc atgactgagt cacctccagg agctcacctg acccccagag aagacctacc    1140
ccaggcagct ccgtgccctg gcttctcccc atgccccaaa tccccccag ccatccctcc    1200
```

```
tggtcctcgt ctacatcaag ggcctcttcc cctcttcctc ccagctctca ggacaggtga    1260
ctgggagacc ttgaaccctc agcctcttcc tttaaaaaaa acaaaacaaa acaaaactgt    1320
gggccattta tttgggattt tggagttgtt tggtttttgt ttttatatct taatagttcg    1380
aaagtaagaa gggagccctg ctatggatgt taagtccaaa ttactcggtt agtgggagca    1440
aaacctatga cttccaaggg gatgaggaga ggttcagagg acaggaggag cctcccccat    1500
tgaaaaaaaa aaaatgggtc aggacattcc ctggatgagg acaatgctag ggtggcatc     1560
tcacatggct gctgctattc ctggtgcttc cccacacttt tgacagatgg agtccttctc    1620
ctaccgcctc ctgccacctc accctacagg cattctctat gtaggaaaca agagccttat    1680
cttatagagt ggggagctga gacacagcct caggtaacac tgacacagct cccgaatgag    1740
gctgggacac tctgcaaacc tctcctcatg gtgctaaggg tggcatgctc ttgacaggaa    1800
acctaaatga ccactcctct catttggaaa gtaatccact gcagtaaaag tttcagacat    1860
gcaagagaga gttttttttt tttttacta caaatttttg ctcccccata aaattatttt     1920
tttattagag ggagtatcca agttttaaaa gtatatagaa ttttttggtt gtaagagaaa    1980
tacatactca ttaggatccc gattaaattc cttgagtaga ctggtgccta ccagaaagca    2040
aagcaaagtt aaacaaaacg aaacaaaatc cttcatatac aaaaagaact ttctgtttgt    2100
attggcagag gtagtgaggt gattcaggta ggctgaaaat cctgggttgc gggagcctca    2160
ctttattcca ttcccacccg ctttgatgtc tatgcttggc tctctgggct gcccctggta    2220
ctgccgaatc ctacacatct cttatcagct ttcctcaaac tttaaggagg ctctgtgagg    2280
gatgggtcat gggaagaccc aagctttccc tccgccagga ttgcaaaagc aagtagactt    2340
ggtctatgca gctcttcttc caacaatttc tttatttgga attagaactt cctttgttag    2400
tatctttgat cttttgactc aagcacattt tggaagggct cccttacaaa agtagaattt    2460
aaaacagagg atacagttaa agagcaaccc aaaggacgct taagaaaccg agaccacttc    2520
accgaacagg actaaggaac actttcgtgc acagaagtca gccgcaatcc aggcacagga    2580
cgaagatggg atacacgtgc tcatctgtct gtcctccttt cctctccctc cccgacgttc    2640
tagttagctt gttgacttgt taaaccttct gttcttaaaa tgaaaagcta gcttacctca    2700
aagaatcttg tttccattcg gaaaccaacg attttgtgtt ttagaatgga cagccctccc    2760
ctcaccactc cctaccttgg cctggtgtcc ttgagacata cggtctttgc ttagtcgtgt    2820
gttggctgct ttgagcagga acgaggcctc caggccctga ggtgggaagg aaggattgga    2880
tgccactgcc ctcctcccca ctttagcatg taggggccag cccatctctt ccagcagggt    2940
cctgctgagt taccatagca accagcaact ccagggtacc acaacagaca atggctcagc    3000
gagccgacgt gtggggatga tgcaggggtt ttggcccagc cagaggaccc agagttgagc    3060
ttcaaatgct agagaagggg agaaacagga tggaagggtg gtttaaggaa ctggcagggg    3120
tctttgagtc acatagagaa gccgttgaag gaggtagggc aggttatctc tgttccagtc    3180
accccccttcc agccccatcc cacttctgtt tcaaactaaa gctcccacct cgaacattga   3240
ccctttgtta gaacaaagca aagcatatct ttaaacaaca gtgttaaaat gagcctcaaa    3300
tgtatgtgga tgagatctct aagaagaggg tcttctggtt ttgattttta aagaagagta    3360
tcctagtaaa atattaaaaa aaaattaaaa agtttttaaa aaggaaacct atgctattta    3420
aattggagcc cagttgtaac ttggtaaagg caagcttctg tacctttgtt ataattaatt    3480
gtatacctgt gtatgtaaat ataaggcatt cctattttgc agttcagaac aaaaaaaact    3540
tatttgtaat atagaataaa gtttattaaa aaataagaaa aaaaaaaaaa aaaaaaa       3597
```

<210> 71
<211> 855
<212> DNA
<213> Homo sapiens

<400> 71

```
cgctcaatta tctactcgag tctagactcg aggcggccgc ccattgtgca ctaaagcagg      60
ggatagcaac ggcgtccctc ctccccgctc agctgcagcc cgcagtcctc acagtggtaa     120
catgccacgt ggtagtctct gtccatggac accacacgga tggttgtctc gcagccctgt     180
gcagggagga taggacgggc acaggaggcg cattttggtg caaaaaccgt gtgatagtct     240
cgcacgcagt agatgttgtt ctccacgtcc acggtgaagg gaaccccgtc caggcactca     300
ttgcacacgg agcaccggaa gcagcctggg tggtaggact tgcccagggc ctgcaggatc     360
atttccatga tgagatgtcc acacacgctg catttgtcgg ccgtctgctg gaacccggag     420
tacaggaagt cctcctggca gtacactttc tcacccacgt tgtagaacgc cttcccacgg     480
agtcgtctcc cacacgagtc gcaggtgaag cagtcagtgt gataaagact ccccattgcc     540
tggcacgcct gctgggctcc gtagatgcca agcccacact tgatgcaaat gccgaagtag     600
```

```
tcccgcgccg tgcgcgcctc gagcgcccgc tccagctccc gggtgagcgc ctccagccgc      660
cgctcggccg cgcttgggcc gccctcccgg ccagggggca gcgggagtgc aggcagcggg      720
aagggggccg gccccgcagg ctccggggag cgagcgggcg cggggcaggc gccgggcggg      780
aggaagtcag cgtaggcttt agtgcacaat gggaattcgg atccggtcga cactagttct      840
aggatctctt ttttt                                                       855

<210> 72
<211> 3791
<212> DNA
<213> Homo sapiens

<400> 72
acagacggcg ggtgaacatg gcgtcctcga cttggtctga gacgtgatag gcctgccttc       60
tggttgaaga tgtggcgagt gaaaaaactg agcctcagcc tgtcgccttc gccccagacg      120
ggaaaaccat ctatgagaac tcctctccgt gaacttaccc tgcagcccgg tgccctcacc      180
acctctggaa aaagatcccc cgcttgctcc tcgctgaccc catcactgtg caagctgggg      240
ctgcaggaag gcagcaacaa ctcgtctcca gtggattttg taaataacaa gaggacagac      300
ttatcttcag aacatttcag tcattcctca aagtggctag aaacttgtca gcatgaatca      360
gatgagcagc ctctagatcc aattccccaa attagctcta ctcctaaaac gtctgaggaa      420
gcagtagacc cactgggcaa ttatatggtt aaaaccatcg tccttgtacc atctccactg      480
gggcagcaac aagacatgat atttgaggcc cgtttagata ccatggcaga gacaaacagc      540
atatctttaa atggaccttt gagaacagac gatctggtga gagaggaggt ggcaccctgc      600
atgggagaca ggttttcaga agttgctgct gtatctgaga aacctatctt tcaggaatct      660
ccgtcccatc tcttagagga gtctccacca aatccctgtt ctgaacaact acattgctcc      720
aaggaaagcc tgagcagtag aactgaggct gtgcgtgagg acttagtacc ttctgaaagt      780
aacgccttct tgccttcctc tgttctctgg ctttcccctt caactgcctt ggcagcagat      840
ttccgtgtca atcatgtgga cccagaggag gaaattgtag agcatggagc tatggaggaa      900
agagaaatga ggtttcccac acatcctaag gagtctgaaa cagaagatca agcacttgtc      960
tcaagtgtgg aagatattct gtccacatgc ctgacaccaa atctagtaga aatggaatcc     1020
caagaagctc caggcccagc agtagaagat gttggtagga ttcttggctc tgatacagag     1080
tcttggatgt ccccactggc ctggctggaa aaaggtgtaa atacctccgt catgctggaa     1140
aatctccgcc aaagcttatc ccttccctcg atgcttcggg atgctgcaat tggcactacc     1200
cctttctcta cttgctcggt ggggacttgg tttactcctt cagcaccaca ggaaaagagt     1260
acaaacacat cccagacagg cctggttggc accaagcaca gtacttctga gacagagcag     1320
ctcctgtgtg gccggcctcc agatctgact gccttgtctc gacatgactt ggaagataac     1380
ctgctgagct ctcttgtcat tgtggagttt ctctcccgcc agcttcggga ctggaagagc     1440
cagctggctg tccctcaccc agaaacccag gacagtagca cacagactga cacatctcac     1500
agtgggataa ctaataaact tcagcatctt aaggagagcc atgagatggg acaggcccta     1560
cagcaggcca gaaatgtcat gcaatcatgg gtgcttatct ctaaagagct gatatccttg     1620
cttcacctat ccctgttgca tttagaagaa gataagacta ctgtgaatca ggagtctcgg     1680
cgtgcagaaa cattggtctg ttgctgtttt gatttgctga agaaattgag ggcaaagctc     1740
cagagcctca aagcagaaag ggaggaggca aggcacagag aggaaatggc tctcagaggc     1800
aaggatgcgg cagagatagt gttggaggct ttctgtgcac acgccagcca gcgcatcagc     1860
cagctggaac aggacctagc atccatgcgg gaattcagag gccttctgaa ggatgcccag     1920
acccaactgg tagggcttca tgccaagcaa gaagagctgg ttcagcagac agtgagtctt     1980
acttctacct tgcaacaaga ctggaggtcc atgcaactgg attatacaac atggacagct     2040
ttgctgagtc ggtcccgaca actcacagag aaactcacag tcaagagcca gcaagccctg     2100
caggaacgtg atgtggcaat tgaggaaaag caggaggttt ctagggtgct ggaacaagtc     2160
tctgcccagt tagaggagtg caaaggccaa acagaacaac tggagttgga aaacattcgt     2220
ctagcaacag atctccgggc tcagttgcag attctggcca acatggacag ccagctaaaa     2280
gagctacaga gtcagcatac ccattgtgcc caggacctgg ctatgaagga tgagttactc     2340
tgccagctta cccagagcaa tgaggagcag gctgctcaat gcgtaaagga agagatggca     2400
ctaaaacaca tgcaggcaga actgcagcag caacaagctg tcctggccaa agaggtgcgg     2460
gacctgaaag agaccttgga gtttgcagac caggagaatc aggttgctca cctggagctg     2520
ggtcaggttg agtgtcaatt gaaaaccaca ctggaagtgc tccgggagcg cagcttgcag     2580
tgtgagaacc tcaaggacac tgtagagaac ctaacggcta aactggccag caccatagca     2640
gataaccagg agcaagatct ggagaaaaca cggcagtact ctcaaaagct agggctgctg     2700
```

actgagcaac tacagagcct gactctcttt ctacagacaa aactaaagga gaagactgaa 2760
caagagaccc ttctgctgag tacagcctgt cctcccaccc aggaacaccc tctgcctaat 2820
gacaggacct tcctgggaag catcttgaca gcagtggcag atgaagagcc agaatcaact 2880
cctgtgccct tgcttggaag tgacaagagt gctttcaccc gagtagcatc aatggtttcc 2940
cttcagcccg cagagacccc aggcatggag gagagcctgg cagaaatgag tattatgact 3000
actgagcttc agagtctttg ttccctgcta caagagtcta aagaagaagc catcaggact 3060
ctgcagcgaa aaatttgtga gctgcaagct aggctgcagg cccaggaaga acagcatcag 3120
gaagtccaga aggcaaaaga agcagacata gagaagctga accaggcctt gtgcttgcgc 3180
tacaagaatg aaaaggagct ccaggaagtg atacagcaga atgagaagat cctagaacag 3240
atagacaaga gtggcgagct cataagcctt agagaggagg tgacccacct tacccgctca 3300
cttcggcgtg cggagacaga gaccaaagtg ctccaggagg cctggcaggc cagctggact 3360
ccaactgcca gcctatggcc accaattgga tccaggagaa agtgtggctc tctcaggagg 3420
tggacaaact gagagtgatg ttcctggaga tgaaaaatga gaaggaaaac tcctgatcaa 3480
gttccagagc ccatagaaat atcctagagg agaaccttcg gcgctctgac aaggagttag 3540
aaaaactaga tgacattgtt cagcatattt ataagaccct gctctctatt ccagaggtgg 3600
tgaggggatg caaagaacta cagggattgc tggaatttct gagctaagaa actgaaagcc 3660
agaatttgtt tcacctcttt ttacctgcaa taccccctta ccccaatacc aagaccaact 3720
ggcatagagc caactgagat aaatgctatt taaataaagt gtatttaatg aaaaaaaaaa 3780
aaaaaaaaaa a 3791

<210> 73
<211> 1683
<212> DNA
<213> Homo sapiens

<400> 73
ctctgagtgt ccagtggtca gttgccccag gatggggacc acagccagag cagccttggt 60
cttgacctat ttggctgttg cttctgctgc ctctgaggga ggcttcacgg ctacaggaca 120
gaggcagctg aggccagagc actttcaaga agttggctac gcagctcccc cctccccacc 180
cctatcccga agcctcccca tggatcaccc tgactcctct cagcatggcc ctccctttga 240
gggacagagt caagtgcagc cccctccctc tcaggaggcc acccctctcc aacaggaaaa 300
gctgctacct gcccaactcc ctgctgaaaa ggaagtgggt ccccctctcc ctcaggaagc 360
tgtccccctc caaaaagagc tgccctctct ccagcacccc aatgaacaga aggaaggaac 420
gccagctcca tttggggacc agagccatcc agaacctgag tcctggaatg cagcccagca 480
ctgccaacag gaccggtccc aagggggctg ggccaccggg ctggatggct tcccccctgg 540
gcggccttct ccagacaatc tgaaccaaat ctgccttcct aaccgtcagc atgtggtata 600
tggtccctgg aacctaccac agtccagcta ctcccacctc actcgccagg gtgagaccct 660
caatttcctg gagattggat attcccgctg ctgccactgc cgcagccaca caaaccgcct 720
agagtgtgcc aaacttgtgt gggaggaagc aatgagccga ttctgtgagg ccgagttctc 780
ggtcaagacc cgaccccact ggtgctgcac gcggcagggg gaggctcggt tctcctgctt 840
ccaggaggaa gctccccagc cacactacca gctccgggcc tgccccagcc atcagcctga 900
tatttcctcg ggtcttgagc tgcctttccc tcctggggtg cccacattgg acaatatcaa 960
gaacatctgc cacctgaggc gcttccgctc tgtgccacgc aacctgccag ctactgaccc 1020
cctacaaagg gagctgctgg cactgatcca gctggagagg gagttccagc gctgctgccg 1080
ccaggggaac aatcacacct gtacatggaa ggcctgggag gataccettg acaaatactg 1140
tgaccgggag tatgctgtga agacccacca ccacttgtgt tgccgccacc ctcccagccc 1200
tactcgggat gagtgctttg cccgtcgggc tccttacccc aactatgacc gggacatctt 1260
gaccattgac atcagtcgag tcacccccaa cctcatgggc cacctctgtg gaaaccaaag 1320
agttctcacc aagcataaac atattcctgg gctgatccac aacatgactg cccgctgctg 1380
tgacctgcca tttccagaac aggcctgctg tgcagaggag gagaaattaa ccttcatcaa 1440
tgatctgtgt ggtccccgac gtaacatctg gcgagaccct gccctctgct gttacctgag 1500
tcctggggat gaacaggtca actgcttcaa catcaattat ctgaggaacg tggctctagt 1560
gtctggagac actgagaacg ccaagggcca gggggagcag ggctcaactg gaggaacaaa 1620
tatcagctcc acctctgagc ccaaggaaga atgagtcacc ccagagccct agagggtcag 1680
atg 1683

<210> 74

<211> 1696
<212> DNA
<213> Homo sapiens

<400> 74
cacctaaaag ccaaaatggg aaaggaaaag actcatatca acattgtcgt cattggacac 60
gtagattcgg gcaagtccac cactactggc catctgatct ataaatgcgg tggcatcgac 120
aaaagaacca ttgaaaaatt tgagaaggag gctgctgaga tgggaaaggg ctccttcaag 180
tatgcctggg tcttggataa actgaaagct gagcgtgaac gtggtatcac cattgatatc 240
tccttgtgga aatttgagac cagcaagtac tatgtgacta tcattgatgc cccaggacac 300
agagacttta tcaaaaacat gattacaggg acatctcagg ctgactgtgc tgtcctgatt 360
gttgctgctg gtgttggtga atttgaagct ggtatctcca agaatgggca gacccgagag 420
catgcccttc tggcttacac actgggtgtg aaacaactaa ttgtcggtgt taacaaaatg 480
gattccactg agccacccta cagccagaag agatatgagg aaattgttaa ggaagtcagc 540
acttacatta agaaaattgg ctacaacccc gacacagtag catttgtgcc aatttctggt 600
tggaatggtg acaacatgct ggagccaagt gctaacatgc cttggttcaa gggatggaaa 660
gtcacccgta aggatggcaa tgccagtgga accacgctgc ttgaggctgt ggactgcatc 720
ctaccaccaa ctcgtccaac tgacaagccc ttgcgcctgc ctctccagga tgtctacaaa 780
attggtggta ttggtactgt tcctgttggc cgagtggaga ctggtgttct caaacccggt 840
atggtggtca cctttgctcc agtcaacgtt acaacggaag taaaatctgt cgaaatgcac 900
catgaagctt tgagtgaagc tcttcctggg gacaatgtgg gcttcaatgt caagaatgtg 960
tctgtcaagg atgttcgtcg tggcaacgtt gctggtgaca gcaaaaatga cccaccaatg 1020
gaagcagctg gcttcactgc tcaggtgatt atcctgaacc atccaggcca aataagcgcc 1080
ggctatgccc ctgtattgga ttgccacacg gctcacattg catgcaagtt tgctgagctg 1140
aaggaaaaga ttgatcgccg ttctggtaaa aagctggaag atggccctaa attcttgaag 1200
tctggtgatg ctgccattgt tgatatggtt cctggcaagc ccatgtgtgt tgagagcttc 1260
tcagactatc cacctttggg tcgctttgct gttcgtgata tgagacagac agttgcggtg 1320
ggtgtcatca aagcagtgga caagaaggct gctggagctg gcaaggtcac caagtctgcc 1380
cagaaagctc agaaggctaa atgaatatta tccctaatac ctgccacccc actcttaatc 1440
agtggtggaa gaacggtctc agaactgttt gtttcaattg gccatttaag tttagtagta 1500
aaagactggt taatgataac aatgcatcgt aaaaccttca gaaggaaagg agaatgtttt 1560
gtggaccact ttggttttct tttttgcgtg tggcagtttt aagttattag tttttaaaat 1620
cagtactttt taatggaaac aacttgacca aaaatttgtc acagaatttt gagacccatt 1680
aaaaaagtta aatgag 1696

<210> 75
<211> 7680
<212> DNA
<213> Homo sapiens

<400> 75
gaagagcaag aggcaggctc agcaaatggt tcagccccag tccccggtgg ctgtcagtca 60
aagcaagccc ggttgttatg acaatggaaa acactatcag ataaatcaac agtgggagcg 120
gacctaccta ggtaatgtgt tggtttgtac ttgttatgga ggaagccgag gttttaactg 180
cgaaagtaaa cctgaagctg aagagacttg ctttgacaag tacactggga acacttaccg 240
agtgggtgac acttatgagc gtcctaaaga ctccatgatc tgggactgta cctgcatcgg 300
ggctgggcga gggagaataa gctgtaccat cgcaaaccgc tgccatgaag gggtcagtc 360
ctacaagatt ggtgacacct ggaggagacc acatgagact ggtggttaca tgttagagtg 420
tgtgtgtctt ggtaatggaa aaggagaatg gacctgcaag cccatagctg agaagtgttt 480
tgatcatgct gctgggactt cctatgtggt cggagaaacg tgggagaagc cctaccaagg 540
ctggatgatg gtagattgta cttgcctggg agaaggcagc ggacgcatca cttgcacttc 600
tagaaataga tgcaacgatc aggacacaag gacatcctat agaattggag acacctggag 660
caagaaggat aatcgaggaa acctgctcca gtgcatctgc acaggcaacg gccgaggaga 720
gtggaagtgt gagaggcaca cctctgtgca gaccacatcg agcggatctg gccccttcac 780
cgatgttcgt gcagctgttt accaaccgca gcctcacccc cagcctcctc cctatggcca 840
ctgtgtcaca gacagtggtg tggtctactc tgtggggatg cagtggttga agacacaagg 900
aaataagcaa atgctttgca cgtgcctggg caacggagtc agctgccaag agacagctgt 960

```
aacccagact tacggtggca acttaaatgg agagccatgt gtcttaccat tcacctacaa 1020
tggcaggacg ttctactcct gcaccacgga agggcgacag gacggacatc tttggtgcag 1080
cacaacttcg aattatgagc aggaccagaa atactctttc tgcacagacc acactgtttt 1140
ggttcagact caaggaggaa attccaatgg tgccttgtgc cacttcccct tcctatacaa 1200
caaccacaat tacactgatt gcacttctga gggcagaaga gacaacatga agtggtgtgg 1260
gaccacacag aactatgatg ccgaccagaa gtttgggttc tgccccatgg ctgcccacga 1320
ggaaatctgc acaaccaatg aaggggtcat gtaccgcatt ggagatcagt gggataagca 1380
gcatgacatg ggtcacatga tgaggtgcac gtgtgttggg aatggtcgtg gggaatggac 1440
atgcattgcc tactcgcaac ttcgagatca gtgcattgtt gatgacatca cttacaatgt 1500
gaacgacaca ttccacaagc gtcatgaaga ggggcacatg ctgaactgta catgcttcgg 1560
tcagggtcgg ggcaggtgga agtgtgatcc cgtcgaccaa tgccaggatt cagagactgg 1620
gacgttttat caaattggag attcatggga gaagtatgtg catggtgtca gataccagtg 1680
ctactgctat ggccgtggca ttggggagtg gcattgccaa cctttacaga cctatccaag 1740
ctcaagtggt cctgtcgaag tatttatcac tgagactccg agtcagccca actcccaccc 1800
catccagtgg aatgcaccac agccatctca catttccaag tacattctca ggtggagacc 1860
taaaaattct gtaggccgtt ggaaggaagc taccatacca ggccacttaa actcctacac 1920
catcaaaggc ctgaagcctg gtgtggtata cgagggccag ctcatcagca tccagcagta 1980
cggccaccaa gaagtgactc gctttgactt caccaccacc agcaccagca cacctgtgac 2040
cagcaacacc gtgacaggag agacgactcc cttttctcct cttgtggcca cttctgaatc 2100
tgtgaccgaa atcacagcca gtagctttgt ggtctcctgg gtctcagctt ccgacaccgt 2160
gtcgggattc cgggtggaat atgagctgag tgaggaggga gatgagccac agtacctgga 2220
tcttccaagc acagccactt ctgtgaacat ccctgacctg cttcctggcc gaaaatacat 2280
tgtaaatgtc tatcagatat ctgaggatgg ggagcagagt ttgatcctgt ctacttcaca 2340
aacaacagcg cctgatgccc ctcctgaccc gactgtggac caagttgatg acacctcaat 2400
tgttgttcgc tggagcagac cccaggctcc catcacaggg tacagaatag tctattcgcc 2460
atcagtagaa ggtagcagca cagaactcaa ccttcctgaa actgcaaact ccgtcaccct 2520
cagtgacttg caacctggtg ttcagtataa catcactatc tatgctgtgg aagaaaatca 2580
agaaagtaca cctgttgtca ttcaacaaga aaccactggc accccacgct cagatacagt 2640
gccctctccc agggacctgc agtttgtgga agtgacagac gtgaaggtca ccatcatgtg 2700
gacaccgcct gagagtgcag tgaccggcta ccgtgtggat gtgatccccg tcaacctgcc 2760
tggcgagcac gggcagaggc tgcccatcag caggaacacc tttgcagaag tcaccgggct 2820
gtcccctggg gtcacctatt acttcaaagt ctttgcagtg agccatggga gggagagcaa 2880
gcctctgact gctcaacaga caaccaaact ggatgctccc actaacctcc agtttgtcaa 2940
tgaaactgat tctactgtcc tggtgagatg gactccacct cgggcccaga taacaggata 3000
ccgactgacc gtgggcctta cccgaagagg ccagcccagg cagtacaatg tgggtccctc 3060
tgtctccaag tacccccctga ggaatctgca gcctgcatct gagtacaccg tatccctcgt 3120
ggccataaag ggcaaccaag agagccccaa agccactgga gtctttacca cactgcagcc 3180
tgggagctct attccacctt acaacaccga ggtgactgag accaccatcg tgatcacatg 3240
gacgcctgct ccaagaattg gttttaagct gggtgtacga ccaagccagg gaggagaggc 3300
accacgagaa gtgacttcag actcaggaag catcgttgtg tccggcttga ctccaggagt 3360
agaatacgtc tacaccatcc aagtcctgag agatggacag gaaagagatg cgccaattgt 3420
aaacaaagtg gtgacaccat tgtctccacc aacaaacttg catctggagg caaaccctga 3480
cactggagtg ctcacagtct cctgggagag gagcaccacc ccagacatta ctggttatag 3540
aattaccaca acccctacaa acggccagca gggaaattct ttggaagaag tggtccatgc 3600
tgatcagagc tcctgcactt ttgataacct gagtcccggc ctggagtaca atgtcagtgt 3660
ttacactgtc aaggatgaca aggaaagtgt ccctatctct gataccatca tcccagctgt 3720
tcctcctccc actgacctgc gattcaccaa cattggtcca gacaccatgc gtgtcacctg 3780
ggctccaccc ccatccattg atttaaccaa cttcctggtg cgttactcac ctgtgaaaaa 3840
tgaggaagat gttgcagagt tgtcaatttc tccttcagac aatgcagtgg tcttaacaaa 3900
tctcctgcct ggtacagaat atgtagtgag tgtctccagt gtctacgaac aacatgagag 3960
cacacctctt agaggaagac agaaaacagg tcttgattcc ccaactggca ttgacttttc 4020
tgatattact gccaactctt ttactgtgca ctggattgct cctcgagcca ccatcactgg 4080
ctacaggatc cgccatcatc ccgagcactt cagtgggaga cctcgagaag atcgggtgcc 4140
ccactctcgg aattccatca ccctcaccaa cctcactcca ggcacagagt atgtggtcag 4200
catcgttgct cttaatggca gagaggaaag tcccttattg attggccaac aatcaacagt 4260
ttctgatgtt ccgagggacc tggaagttgt tgctgcgacc cccaccagcc tactgatcag 4320
ctgggatgct cctgctgtca cagtgagata ttacaggatc acttacggag aaacaggagg 4380
```

```
aaatagccct gtccaggagt tcactgtgcc tgggagcaag tctacagcta ccatcagcgg    4440
ccttaaacct ggagttgatt ataccatcac tgtgtatgct gtcactggcc gtggagacag    4500
ccccgcaagc agcaagccaa tttccattaa ttaccgaaca gaaattgaca aaccatccca    4560
gatgcaagtg accgatgttc aggacaacag cattagtgtc aagtggctgc cttcaagttc    4620
ccctgttact ggttacagag taaccaccac tcccaaaaat ggaccaggac caacaaaaac    4680
taaaactgca ggtccagatc aaacagaaat gactattgaa ggcttgcagc ccacagtgga    4740
gtatgtggtt agtgtctatg ctcagaatcc aagcggagag agtcagcctc tggttcagac    4800
tgcagtaacc aacattgatc gccctaaagg actggcattc actgatgtgg atgtcgattc    4860
catcaaaatt gcttgggaaa gcccacaggg gcaagtttcc aggtacaggg tgacctactc    4920
gagccctgag gatggaatcc atgagctatt ccctgcacct gatggtgaag aagacactgc    4980
agagctgcaa ggcctcagac cgggttctga gtacacagtc agtgtggttg ccttgcacga    5040
tgatatggag agccagcccc tgattggaac ccagtccaca gctattcctg caccaactga    5100
cctgaagttc actcaggtca cacccacaag cctgagcgcc cagtggacac cacccaatgt    5160
tcagctcact ggatatcgag tgcgggtgac ccccaaggag aagaccggac caatgaaaga    5220
aatcaacctt gctcctgaca gctcatccgt ggttgtatca ggacttatgg tggccaccaa    5280
atatgaagtg agtgtctatg ctcttaagga cactttgaca agcagaccag ctcagggtgt    5340
tgtcaccact ctggagaatg tcagcccacc aagaagggct cgtgtgacag atgctactga    5400
gaccaccatc accattagct ggagaaccaa gactgagacg atcactggct tccaagttga    5460
tgccgttcca gccaatggcc agactccaat ccagagaacc atcaagccag atgtcagaag    5520
ctacaccatc acaggtttac aaccaggcac tgactacaag atctacctgt acaccttgaa    5580
tgacaatgct cggagctccc ctgtggtcat cgacgcctcc actgccattg atgcaccatc    5640
caacctgcgt ttcctggcca ccacacccaa ttccttgctg gtatcatggc agccgccacg    5700
tgccaggatt accggctaca tcatcaagta tgagaagcct gggtctcctc ccagagaagt    5760
ggtccctcgg ccccgccctg gtgtcacaga ggctactatt actggcctgg aaccgggaac    5820
cgaatataca atttatgtca ttgccctgaa gaataatcag aagagcgagc ccctgattgg    5880
aaggaaaaag acagacgagc ttccccaact ggtaaccctt ccacacccca atcttcatgg    5940
accagagatc ttggatgttc cttccacagt tcaaaagacc cctttcgtca cccaccctgg    6000
gtatgacact ggaaatggta ttcagcttcc tggcacttct ggtcagcaac ccagtgttgg    6060
gcaacaaatg atctttgagg aacatggttt taggcggacc acaccgccca caacggccac    6120
ccccataagg cataggccaa gaccataccc gccgaatgta ggacaagaag ctctctctca    6180
gacaaccatc tcatgggccc cattccagga cacttctgag tacatcattt catgtcatcc    6240
tgttggcact gatgaagaac ccttacagtt cagggttcct ggaacttcta ccagtgccac    6300
tctgacaggc ctcaccagag gtgccaccta caacatcata gtggaggcac tgaaagacca    6360
gcagaggcat aaggttcggg aagaggttgt taccgtgggc aactctgtca acgaaggctt    6420
gaaccaacct acggatgact cgtgctttga cccctacaca gtttcccatt atgccgttgg    6480
agatgagtgg gaacgaatgt ctgaatcagg ctttaaactg ttgtgccagt gcttaggctt    6540
tggaagtggt catttcagat gtgattcatc tagatggtgc catgacaatg gtgtgaacta    6600
caagattgga gagaagtggg accgtcaggg agaaaatggc cagatgatga gctgcacatg    6660
tcttgggaac ggaaaaggag aattcaagtg tgaccctcat gaggcaacgt gttacgatga    6720
tgggaagaca taccacgtag gagaacagtg gcagaaggaa tatctcggtg ccatttgctc    6780
ctgcacatgc tttggaggcc agcggggctg gcgctgtgac aactgccgca gacctggggg    6840
tgaacccagt cccgaaggca ctactggcca gtcctacaac cagtattctc agagatacca    6900
tcagagaaca aacactaatg ttaattgccc aattgagtgc ttcatgcctt tagatgtaca    6960
ggctgacaga gaagattccc gagagtaaat catctttcca atccagagga acaagcatgt    7020
ctctctgcca agatccatct aaactggagt gatgttagca gacccagctt agagttcttc    7080
tttctttctt aagccctttg ctctggagga agttctccag cttcagctca actcacagct    7140
tctccaagca tcaccctggg agtttcctga gggttttctc ataaatgagg gctgcacatt    7200
gcctgttctg cttcgaagta ttcaataccg ctcagtattt taaatgaagt gattctaaga    7260
tttggtttgg gatcaatagg aaagcatatg cagccaacca agatgcaaat gttttgaaat    7320
gatatgacca aaattttaag taggaaagtc acccaaacac ttctgctttc acttaagtgt    7380
ctggcccgca atactgtagg aacaagcatg atcttgttac tgtgatattt taaatatcca    7440
cagtactcac tttttccaaa tgatcctagt aattgcctag aaatatcttt ctcttacctg    7500
ttatttatca atttttccca gtatttttat acggaaaaaa ttgtattgaa aacacttagt    7560
atgcagttga taagaggaat ttggtataat tatggtgggt gattattttt tatactgtat    7620
gtgccaaagc tttactactg tggaaagaca actgttttaa taaaagattt acattccaca    7680
```

<210> 76
<211> 1316
<212> DNA
<213> Homo sapiens

<400> 76
tcctaatacg actcactata gggctcgagc ggccgcccgg gcaggtcgaa tgcaggcgac 60
ttgcgagctg ggagcgattt aaaacgcttt ggattccccc ggcctgggtg gggagagcga 120
gctgggtgcc ccctagattc cccgcccccg cacctcatga gccgaccctc ggctccatgg 180
agcccggcaa ttatgccacc ttggatggag ccaaggatat cgaaggcttg ctgggagcgg 240
gaggggggcg gaatctggtc gcccactccc ctctgaccag ccacccagcg gcgcctacgc 300
tgatgcctgc tgtcaactat gccccccttgg atctgccagg ctcggcggag ccgccaaagc 360
aatgccaccc atgccctggg gtgccccagg ggacgtcccc agctcccgtg ccttatggtt 420
actttggagg cgggtactac tcctgccgag tgtcccggag ctcgctgaaa ccctgtgccc 480
aggcagccac cctggccgcg taccccgcgg agactcccac ggccggggaa gagtacccca 540
gtcgccccac tgagtttgcc ttctatccgg gatatccggg aacctaccac gctatggcca 600
gttacctgga cgtgtctgtg gtgcagactc tgggtgctcc tggagaaccg cgacatgact 660
ccctgttgcc tgtggacagt taccagtctt gggctctcgc tggtggctgg aacagccaga 720
tgtgttgcca gggagaacag aacccaccag gtccctttttg gaaggcagca tttgcagact 780
ccagcgggca gcaccctcct gacgcctgcg cctttcgtcg cggccgcaag aaacgcattc 840
cgtacagcaa ggggcagttg cgggagctgg agcgggagta tgcggctaac aagttcatca 900
ccaaggacaa gaggcgcaag atctcggcag ccaccagcct ctcggagcgc cagattacca 960
tctggtttca gaaccgccgg gtcaaagaga agaaggttct cgccaaggtg aagaacagcg 1020
ctacccctta agagatctcc ttgcctgggt gggaggagcg aaagtggggg tgtcctgggg 1080
agaccagaaa cctgccaagc ccaggctggg gccaaggact ctgctgagag gcccctagag 1140
acaacaccct tcccaggcca ctggctgctg gactgttcct caggagcggc ctgggtaccc 1200
agtatgtgca gggagacgga accccatgtg acaggcccac tccaccaggg ttcccaaaga 1260
acctggccca gtcataatca ttcatcctca cagtggcaat aatcacgata accagt 1316

<210> 77
<211> 566
<212> DNA
<213> Homo sapiens

<400> 77
cccaccaaac ccataaagag ggtgggtcga cccacgcgtc cgcggacgcg tgggaaatta 60
ttgaattgga aacagaaata gaaaagttta aagctgagaa cgcatcttta gctaaacttc 120
gcattgaacg agaaagtgcc ttggaaaaac tcaggaaaga aattgcagac ttcgaacaac 180
agaaagcaaa agaattagct cgaatagaag agtttaaaaa ggaggagatg aggaagctac 240
aaaaggaacg taaagttttt gaaaagtata ctacagctgc aagaactttt ccagataaaa 300
aggaacgtga agaaatacag actttaaaac agcaaatagc agatttacgg gaagatttga 360
aaagaaagga gaccaaatgg tcaagtacac acagccgtct cagaagccag atacaaatgt 420
tagtcagaga gaacacagac ctccgggaag aaataaaagt gatggaaaga ttccgactgg 480
atgcctggaa gagagcagaa gccatagaga gcagcctcga ggtggagaag aaggacaagc 540
ttgcgaacac atctgttcga tttcaa 566

<210> 78
<211> 5067
<212> DNA
<213> Homo sapiens

<400> 78
gcccggacac ctgtctgcag catggataag tatgacgacc tgggcctgga ggccagtaaa 60
ttcatcgagg acctgaacat gtatgaggcc tctaaggatg ggctcttccg agtggacaag 120
ggtgcaggca acaaccccga gtttgaggaa actcgcaggg tgttcgccac caagatggcc 180
aaaatccacc tccagcagca gcagcagcag ctcctgcagg aggagactct gcccaggggg 240
agtagaggcc ctgtcaatgg agggggccgc ctgggcccac aggcccgttg ggaagttgtg 300

```
ggcagcaagc tgactgtgga tggtgctgcc aagcctcctc ttgctgcctc gacaggggca     360
cctggggcag tcaccaccct cgctgctggg cagccccgt acccaccgca ggagcagaga     420
tccaggccat acctgcatgg cacgaggcat ggcagccagg actgtggttc cagggagagc     480
ctggcgactt ctgagatgtc tgctttccac cagccaggcc cctgtgagga tccttcctgc     540
ctcactcatg gagactatta tgacaacctc tccttggcaa gcccaaagtg gggtgacaaa     600
ccaggagtgt cccccagcat cggcctgagt gtagggagtg ggtggcctag ctccccgggg     660
agtgacccac cactgcccaa accctgcggg gaccatcccc taaatcaccg acagctctcc     720
ctgagctcca gcaggtcttc tgagggtagc ctcggtggtc agaatagtgg cattggtggc     780
cgcagcagcg agaagccaac aggcctttgg tccactgcct cctcccagcg ggtgagccct     840
ggcctgcctt ccccaaactt ggagaacgga gcaccagctg tggggcctgt tcagcccagg     900
actccttctg tgtcagcacc cttggccctg agctgcccca ggcaaggagg tcttccaaga     960
tcaaactcgg ggctgggggg tgaggtttca ggtgtgatgt ccaaacccaa tgtggacccc    1020
caaccctggt tccaggatgg gcccaaatct tacctttcca gttctgcccc gtcatcctcg    1080
ccagctggcc tggacggttc acagcagggt gcggtccctg ggctggggcc gaagcctggc    1140
tgcacagacc ttggcactgg tcccaagctc agccccacca gtcttgtcca tccagtgatg    1200
tccaccctgc ctgagttatc ttgtaaagag ggtcccctgg gctggtcttc tgatggtagc    1260
ctgggatctg tgctcctgga cagccccagc tcccctaggg taaggctgcc ctgccagccc    1320
ctcgtcccag gtcctgagct gagaccctct gctgctgagt tgaaattaga agccctcacc    1380
caacgtctgg agcgagagat ggatgctcac ccgaaggctg attactttgg agcctgtgtg    1440
aaatgcagca aaggggtgtt tggggctggc caggcctgtc aggccatggg gaacctctac    1500
catgacacat gcttcacctg tgcagcttgc agccggaagc tgagaggaaa agccttttat    1560
tttgtcaacg gcaaagtgtt ttgtgaagaa gacttcctgt actctggttt ccagcagtcg    1620
gctgacaggt gttttctttg tggacatctg atcatggaca tgatcctgca agccctgggg    1680
aagtcctacc accccggctg tttccgctgt gtcatctgta atgagtgttt ggatggggtg    1740
cccttcaccg tggactcaga gaacaagatc tactgtgtcc gagattacca caaggtgctg    1800
gcccccaagt gtgcagcctg tgggcttccc atccttccac ctgagggctc agatgagacc    1860
atccgtgtcg tgtccatgga cagagactac cacgtggagt gttaccactg cgaggactgt    1920
ggtctggagc tcaatgatga agatggccac cgctgttatc cgctggagga ccacctgttc    1980
tgtcactcct gccacgtgaa gaggctggag aagagaccct catctacagc ccttcaccag    2040
caccacttct agccagagcc acttgcagac atcacggcag gggatgagga gccggggttg    2100
ctgctgctgc ttccggtggc cctggggtg gaagtggggt aggggaagag gaggggcagg     2160
agggagagtt cctgtgagca tgtgggggt gcctttcctt taaccaggga ggtgaacact     2220
acctgcctcc tgcgtgtatt ttccaagtgc ttttctctgt tgccacattt tcctcaggtt    2280
actcaggaaa atgctccagc atgtgcgagc acatgacctg aggttgcatc atagcaccaa    2340
aggaatcctc ctgtcccctc tgggaacatt tcatgcttca gagggagagg tttttattga    2400
gcttgtttca caatatcccc ttgaagggac agctcagctg ccaatacatt caaccctttc    2460
tcttccttca ggaaaatacc tatacccaaa tgttccctcc cccgacatat atcatggcat    2520
gacttaaggc ttcttttcac ctgagagctt cagttcttct gcagaatggc tgcaaattta    2580
attgcattaa ggcaagaagg aagctctaat gtgtgctttg tatcctaaga taaatttgct    2640
tagaaaacca gagtcaagat ttgaaatagg tgaggcaggg tttcctcctt agacactgac    2700
agcattctcc gtaccccttc aaatccttac tctcctaaag gcagctgagt ccgcgacaga    2760
aatttgccct atgggagtaa aacatacttt gggagaagaa cttggtgcag gcaccaggat    2820
ttttttttt gcccacgtgt ttgcgctgtt tttctctgga gttctcaaga gttggtgact     2880
tggaaggccg cttctgcaag gcaagtctca ggaacccatg caggtacatc gcttgcacct    2940
gtttttagct tatttaatga cgggcttttg ggaagagctg cccgcatact gagagacagc    3000
ttcttataaa caaggagagt ttttgtgtgt gcgagatctc taagccagcg tgggagggag    3060
cgcctcagga taagttatta tattcatttc gttggtttct ctcctgccca attcttggca    3120
caggcattat gtttgaagaa accaggataa ggtacactgc ttttgtctgt ttaatttttt    3180
tagttgtttc ccttcacttt cagtcttcca cacacaaaaa atacctcaca gagcttcacc    3240
aaatcacaga ttcaggagga atttggcttt cacactggac tcagatacct tcttcagtgt    3300
gttggaaatc actggcttca cacaggccca actccaactg gtcagggcag agtgatcgta    3360
actaaaggtc agtggggaat agatccgatt cagtgctttt gccttatgca tttcagcatc    3420
ctggctcccc agggtggcag gagctgagga agggccacac actggcaaga tttcaagacc    3480
actctctgca ctgaagaggt aaaatttgca ctgcaagtca catccctgag gccagaggtc    3540
agtacccttt ggtatttcga ttagaagaag ctgcaaaaga aaggcagccc attttaccat    3600
tgccagccag gccggggaca caggagccgg tgtgtgcact ctgcctccta acattgcacc    3660
cagagcaaga ggactgggtg ctgggctgca gaggccggtc agtggagccc ctagcacgtg    3720
```

tgaactcagg cttttcattg ggcccggctc cacttctagg ccatgttttg actcatttgg 3780
taaccattgc ctgtaagcag cacagaattg gtgccatgga ttatcttttc catgttgatg 3840
gaattcattc tgttggaatc ctttggccag atgtcacttc agccagggtg tgcatcatca 3900
ttggttcttt ttcacaggct gagcctcctg aaaacccatg aacgctgggg ctggggaagt 3960
gaaccctgag gtggggaccc tctcttccca tcaaatcatc cagctcagtg tggggcgtgg 4020
caggggggta aatgaagcca gccaatgtgt taacctgtct ctgtcaacct aagaatgttg 4080
gccttactga cacacctttg ctccatgttc aagaccagaa gtagctggga tttgtttgca 4140
aattgggtaa ttagtttaaa aatctgtgat tacattttta aatgaaattt tcaaagtggc 4200
ctagattgag gtgattcaga taggtttgcg aatataccat tttatattgt tgagaaagaa 4260
caaaaaggga atttccagat gtcctagaaa tcctagcaac agatttctct ggttgtcagt 4320
ttccctggag aaggcgccag ataggaatct ccaatcagtt gtttttctct tcgcttcagg 4380
cccttacaca aaagccatga agagatgttc acctacccgg tattttaaat gttctgtaaa 4440
ttattagcca aatagaactg taatggggtt gtatttatgg gcgcctagaa agaaaacaca 4500
aggacttggt aggccaggaa gaaaagattt taaaatttag aatgaatagc ccttctgggt 4560
tttctttttg acaattcttg gacttgaggt aaaacaagga ggattgtggc cggatttcag 4620
atcccaaagc cagcctccat cttaggcctt tgcctcattg tgccttttag gttttcttac 4680
ccaccgtctc ctgttttgtc tttttttct tttctcctac ccctatcttg ggacattcag 4740
aaactgcctg ggtggtttga gaagagacaa cccagtttga tctgcaatac aaggatccat 4800
tcgtaatctc tctctcactg atgttattcc cccatctgcc gtcttggttc atctcaccac 4860
agaagggcat ttagtcctac ccagccatcg gctgcggtat gacagcagga tggcacttcc 4920
catttctctg tggttagtgc tcgagtgaaa acctctttca gctgagtcct ctgaggttct 4980
gctgttgagt cctgggtggc tgatggaatg attgaggagg tctggtcacc ctcaagcgcc 5040
gtcatcgcct tgtttccatg ggcttct 5067

<210> 79
<211> 950
<212> DNA
<213> Homo sapiens

<400> 79
tcgaccggat ccgaattccc attgtgcact aaagcgtctc cctgctccgc ggcccgggct 60
ggcgggcggg cgctcggctg gcggctgcag cagcagaggg agacccgcgg caaccccggc 120
aacccagggc tcggcgtcgc tgccaccatg acgggaagca atatgtcgga cgccttggcc 180
aacgccgtgt gccagcgctg ccaggcccgc ttctccccgg ccgagcgcat tgtcaacagc 240
aatggggagc tgtaccatga gcactgcttc gtgtgtgccc agtgcttccg gcccttcccc 300
gaggggctct tctatgagtt tgaaggccgg aagtactgcg aacacgactt ccaaatgctg 360
tttgctccgt gctgtggatc ctgcggtgag ttcatcattg gccgcgtcat caaggccatg 420
aacaacaact ggcacccggg ctgcttccgc tgcgagctgt gtgatgtgga gctggctgac 480
ctgggctttg tgaagaatgc cggcaggcat ctctgccggc cttgccacaa ccgtgagaag 540
gccaagggcc tgggcaagta catctgccag cggtgccacc tggtcatcga cgagcagccc 600
ctcatgttca ggagcgacgc ctaccaccct gaccacttca actgcaccca ctgtgggaag 660
gagctgacag ccgaggcccg cgagctgaag ggtgagctct actgcctgcc ctgccatgac 720
aagatgggcg ttcccatctg cggggcctgc cgccggccca tcgagggccg agtggtcaac 780
gcgctgggca agcagtggca cgtggagcac tttgtctgtg ccaagtgtga gaagccattc 840
ctggggcacc ggcactatga gaagaagggc ctggcctact gcgagcttta gtgcacaatg 900
ggcggccgcc tcgagtctag actcgagtag ataattgagc ggaatttctt 950

<210> 80
<211> 2346
<212> DNA
<213> Homo sapiens

<400> 80
ccgccgtcgc ccccgcctcc ccctgcctca gcggctgccc ccgccagcgg gccgcccgct 60
cccccgggcc ttgcagcggg ccccggcccg gctggagggg ccccgacccc agctctggtg 120
gcgggcagca gcgccgcggc cccctccct cacggggact cggccctgaa cgagcaggag 180

```
aaggagttgc agcggcggct gaagcgtctc tacccggccg tggacgaaca agagacgccg      240
ctgcctcggt cctggagccc gaaggacaag ttcagctaca tcggcctctc tcagaacaac      300
ctgcgggtgc actacaaagg tcatggcaaa accccaaaag atgccgcgtc agttcgagcc      360
acgcatccaa taccagcagc ctgtgggatt tattatttg aagtaaaaat tgtcagtaag       420
ggaagagatg gttacatggg aattggtctt tctgctcaag gtgtgaacat gaatagacta      480
ccaggttggg ataagcattc atatggttac catggggatg atggacattc gttttgttct      540
tctggaactg gacaacctta tggaccaact ttcactactg gtgatgtcat tggctgttgt      600
gttaatctta tcaacaatac ctgcttttac accaagaatg gacatagttt aggtattgct      660
ttcactgacc taccgccaaa tttgtatcct actgtggggc ttcaaacacc aggagaagtg      720
gtcgatgcca attttgggca acatcctttc gtgtttgata tagaagacta tatgcgggag      780
tggagaacca aaatccaggc acagatagat cgatttccta tcggagatcg agaaggagaa      840
tggcagacca tgatacaaaa aatggtttca tcttatttag tccaccatgg gtactgtgcc      900
acagcagagg cctttgccag atctacagac cagaccgttc tagaagaatt agcttccatt      960
aagaatagac aaagaattca gaaattggta ttagcaggaa gaatgggaga agccattgaa     1020
acaacacaac agttataccc aagtttactt gaaagaaatc ctaatctcct tttcacatta     1080
aaagtgcgtc agtttataga aatggtgaat ggtacagata gtgaagtacg atgtttggga     1140
ggccgaagtc caaagtctca agacagttat cctgttagtc ctcgaccttt tagtagtcca     1200
agtatgagcc ccagccatgg aatgaatatc cacaatttag catcaggcaa aggaagcacc     1260
gcacattttt caggttttga aagttgtagt aatggtgtaa tatcaaataa agcacatcaa     1320
tcatattgcc atagtaataa acaccagtca tccaacttga atgtaccaga actaaacagt     1380
ataaatatgt caagatcaca gcaagttaat aacttcacca gtaatgatgt agacatggaa     1440
acagatcact actccaatgg agttggagaa acttcatcca atggtttcct aaatggtagc     1500
tctaaacatg accacgaaat ggaagattgt gacaccgaaa tggaagttga ttcaagtcag     1560
ttgagacgcc agttgtgtgg aggaagtcag gccgccatag aaagaatgat ccactttgga     1620
cgagagctgc aagcaatgag tgaacagcta aggagagact gtggcaagaa cactgcaaac     1680
aaaaaaatgt tgaaggatgc attcagtcta ctagcatatt cagatccctg gaacagccca     1740
gttggaaatc agcttgaccc gattcagaga gaacctgtgt gctcagctct taacagtgca     1800
atattagaaa cccacaatct gccaaagcaa cctccacttg ccctagcaat gggacaggcc     1860
acacaatgtc taggactgat ggctcgatca ggaattggat cctgcgcatt tgccacagtg     1920
gaagactacc tacattagct atgcatttca agagctcaca cttatattgt ggcatatagt     1980
caacatggaa gtagaccagc tctgctgatt tgaaatttag atttttaaa ttatgtactg      2040
gggacaggtt tttgtcgctt tacattgctt cctagtttac agcatgatgc aaatgatttt     2100
ctaacttagt gttaggagaa attattttcc atctttaacc tcttagttgt ctaagagtta     2160
aatattactg aatttcagac gttcaaattg atcatcacaa atcctttaaa acaattacct     2220
aaaagaaacc aaaaatcctg ccttctttgt gggggagggg ggagagaggg gaaggaaatg     2280
gaacaagttg tgtttgtgtt agcatgtggg tgatgtaaac ttcaaattgg gagatgttcc     2340
gacccc                                                                2346
```

<210> 81
<211> 2512
<212> DNA
<213> Homo sapiens

<400> 81

```
caatgcactg acggatatga gtgggatcct gtgagacagc aatgcaaaga tattgatgaa       60
tgtgacattg tcccagacgc ttgtaaaggt ggaatgaagt gtgtcaacca ctatggagga      120
tacctctgcc ttccgaaaac agcccagatt attgtcaata atgaacagcc tcagcaggaa      180
acacaaccag cagaaggaac ctcaggggca accaccgggg ttgtagctgc cagcagcatg      240
gcaaccagtg gagtgttgcc cgggggtggt tttgtggcca gtgctgctgc agtcgcaggc      300
cctgaaatgc agactggccg aaataacttt gtcatccggc ggaacccagc tgaccctcag      360
cgcattccct ccaacccttc ccaccgtatc cagtgtgcag caggctacga gcaaagtgaa      420
cacaacgtgt gccaagacat agacgagtgc actgcaggga cgcacaactg tagagcagac      480
caagtgtgca tcaatttacg ggatcctttt gcatgtcagt gccctcctgg atatcagaag      540
cgaggggagc agtgcgtaga catagatgaa tgtaccatcc ctccatattg ccaccaaaga      600
tgcgtgaata caccaggctc atttatttgc cagtgcagtc ctgggtttca attggcagca      660
aacaactata cctgcgtaga tataaatgaa tgtgatgcca gcaatcaatg tgctcagcag      720
tgctacaaca ttcttggttc attcatctgt cagtgcaatc aaggatatga gctaagcagt      780
```

gacaggctca actgtgaaga cattgatgaa tgcagaacct caagctacct gtgtcaatat 840
caatgtgtca atgaacctgg gaaattctca tgtatgtgcc cccagggata ccaagtggtg 900
agaagtagaa catgtcaaga tataaatgag tgtgagacca caaatgaatg ccgggaggat 960
gaaatgtgtt ggaattatca tggcggcttc cgttgttatc cacgaaatcc ttgtcaagat 1020
ccctacattc taacaccaga gaaccgatgt gtttgcccag tctcaaatgc catgtgccga 1080
gaactgcccc agtcaatagt ctacaaatac atgagcatcc gatctgatag gtctgtgcca 1140
tcagacatct tccagataca ggccacaact atttatgcca acaccatcaa tacttttcgg 1200
attaaatctg gaaatgaaaa tggagagttc tacctacgac aaacaagtcc tgtaagtgca 1260
atgcttgtgc tcgtgaagtc attatcagga ccaagagaac atatcgtgga cctggagatg 1320
ctgacagtca gcagtatagg gaccttccgc acaagctctg tgttaagatt gacaataata 1380
gtggggccat tttcatttta gtcttttcta agagtcaacc acaggcattt aagtcagcca 1440
aagaatattg ttaccttaaa gcactatttt atttatagat atatctagtg catctacatc 1500
tctatactgt acactcaccc ataacaaaca attacaccat ggtataaagt gggcatttaa 1560
tatgtaaaga ttcaaagttt gtctttatta ctatatgtaa attagacatt aatccactaa 1620
actggtcttc ttcaagagag ctaagtatac actatctggt gaaacttgga ttctttccta 1680
taaaagtggg accaagcaat gatgatcttc tgtggtgctt aaggaaactt actagagctc 1740
cactaacagt ctcataagga ggcagccatc ataaccattg aatagcatgc aagggtaaga 1800
atgagttttt aactgctttg taagaaaatg gaaaaggtca ataaagatat atttctttag 1860
aaaatgggga tctgccatat ttgtgttggt ttttattttc atatccagcc taaaggtggt 1920
tgtttattat atagtaataa atcattgctg tacaacatgc tggtttctgt agggtatttt 1980
taattttgtc agaaatttta gattgtgaat attttgtaaa aaacagtaag caaaattttc 2040
cagaattccc aaaatgaacc agataccccc tagaaaatta tactattgag aaatctatgg 2100
ggaggatatg agaaaataaa ttccttctaa accacattgg aactgacctg aagaagcaaa 2160
ctcggaaaat ataataacat ccctgaattc aggcattcac aagatgcaga acaaaatgga 2220
taaaaggtat ttcactggag aagttttaat ttctaagtaa aatttaaatc ctaacacttc 2280
actaatttat aactaaaatt tctcatcttc gtacttgatg ctcacagagg aagaaaatga 2340
tgatggtttt tattcctggc atccagagtg acagtgaact taagcaaatt accctcctac 2400
ccaattctat ggaatatttt atacgtctcc ttgtttaaaa tctgactgct ttactttgat 2460
gtatcatatt tttaaataaa aataaatatt cctttagaag atcactctaa aa 2512

<210> 82
<211> 2306
<212> DNA
<213> Homo sapiens

<400> 82
gggcgggagc tgcacgcgcc gtggctccgg atctcttcgt ctttgcagcg tacgcccgag 60
tcggtcagcg ccggaggacc tcagcagcca tgtcgaagcc ccatagtgaa gccgggactg 120
ccttcattca gacccagcag ctgcacgcag ccatggctga cacattcctg gagcacatgt 180
gccgcctgga cattgattca ccacccatca cagcccggaa cactggcatc atctgtacca 240
ttggcccagc ttcccgatca gtggagacgt tgaaggagat gattaagtct ggaatgaatg 300
tggctcgtct gaacttctct catggaactc atgagtacca tgcggagacc atcaagaatg 360
tgcgcacagc cacggaaagc tttgcttctg acccctacct ctaccggccc gttgctgtgg 420
ctctagacac taaaggacct gagatccgaa ctgggctcat caagggcagc ggcactgcag 480
agctggagct gaagaaggga gccactctca aaatcacgct ggataacgcc tacatggaaa 540
agtgtgacga gaacatcctg tggctggact acaagaacat ctgcaaggtg gtggaagtgg 600
gcagcaagat ctacgtggat gatgggctta tttctctcca ggtgaagcag aaaggtgccg 660
acttcctggt gacggaggtg gaaaatggtg gctccttggg cagcaagaag ggtgtgaacc 720
ttcctggggc tgctgtggac ttgcctgctg tgtcggagaa ggacatccag gatctgaagt 780
ttggggtcga gcaggatgtt gatatggtgt ttgcgtcatt catccgcaag gcatctgatg 840
tccatgaagt taggaaggtc ctgggagaga agggaaagaa catcaagatt atcagcaaaa 900
tcgagaatca tgaggggggtt cggaggtttg atgaaatcct ggaggccagt gatgggatca 960
tggtggctcg tggtgatcta ggcattgaga ttcctgcaga gaaggtcttc cttgctcaga 1020
agatgatgat tggacggtgc aaccgagctg ggaagcctgt catctgtgct actcagatgc 1080
tggagagcat gatcaagaag ccccgcccca ctcgggctga aggcagtgat gtggccaatg 1140
cagtcctgga tggagccgac tgcatcatgc tgtctggaga aacagccaaa ggggactatc 1200
ctctggaggc tgtgcgcatg cagcacctga ttgcccgtga ggcagaggct gccatctacc 1260

```
acttgcaatt atttgaggaa ctccgccgcc tggcgcccat taccagcgac cccacagaag    1320
ccaccgccgt gggtgccgtg gaggcctcct tcaagtgctg cagtggggcc ataatcgtcc    1380
tcaccaagtc tggcaggtct gctcaccagg tggccagata ccgcccacgt gcccccatca    1440
ttgctgtgac ccggaatccc cagacagctc gtcaggccca cctgtaccgt ggcatcttcc    1500
ctgtgctgtg caaggaccca gtccaggagg cctgggctga ggacgtggac ctccgggtga    1560
actttgccat gaatgttggc aaggcccgag gcttcttcaa gaagggagat gtggtcattg    1620
tgctgaccgg atggcgccct ggctccggct tcaccaacac catgcgtgtt gttcctgtgc    1680
cgtgatggac cccagagccc ctcctccagc ccctgtccca ccccccttccc ccagcccatc    1740
cattaggcca gcaacgcttg tagaactcac tctgggctgt aacgtggcac tggtaggttg    1800
ggacaccagg gaagaagatc aacgcctcac tgaaacatgg ctgtgtttgc agcctgctct    1860
agtgggacag cccagagcct ggctgcccat catgtggccc cacccaatca agggaagaag    1920
gaggaatgct ggactggagg cccctggagc cagatggcaa gagggtgaca gcttcctttc    1980
ctgtgtgtac tctgtccagt tcctttagaa aaaatggatg cccagaggac tcccaaccct    2040
ggcttggggt caagaaacag ccagcaagag ttagggggcct tagggcactg ggctgttgtt    2100
ccattgaagc cgactctggc cctggcccctt acttgcttct ctagctctct aggcctctcc    2160
agtttgcacc tgtccccacc ctccactcag ctgtcctgca gcaaacactc caccctccac    2220
cttccatttt cccccactac tgcagcacct ccaggcctgt tgctatagag cctacctgta    2280
tgtcaataaa caacagctga agcacc                                        2306
```

```
<210> 83
<211> 2656
<212> DNA
<213> Homo sapiens

<400> 83
gaattcgcgg ccgcagagtc cccgggccaa gatggctgcg cggtgctcca cacgctggtt      60
gctggtggtt gtggggaccc cgcggctgcc ggctatatcg ggtagagggg cccggccgcc     120
cagggagggc gtggtggggg catggctgag ccgcaagctg agcgtccccg cctttgcgtc     180
ttccctgacc tcttgcggcc cccgagcgct gctgacattg agacctggtg tcagccttac     240
aggaacaaaa cataaccctt tcatttgtac tgcctccttc cacacgagtg ccccctttggc     300
caaagaagat tattatcaga tattaggagt gcctcgaaat gccagccaga aagagatcaa     360
gaaagcctat tatcagcttg ccaagaagta tcaccctgac acaaataagg atgatcccaa     420
agccaaggag aagttctccc agctggcaga agcctatgag gttttgagtg atgaggtgaa     480
gaggaagcag tacgatgcct acggctctgc aggcttcgat cctggggcca gcggctccca     540
gcatagctac tggaagggag gccccactgt ggaccccgag gagctgttca ggaagatctt     600
tggcgagttc tcatcctctt catttggaga tttccagacc gtgtttgatc agcctcagga     660
atacttcatg gagttgacat tcaatcaagc tgcaaagggg gtcaacaagg agttcaccgt     720
gaacatcatg gacacgtgtg agcgctgcaa cggcaagggg aacgagcccg gcaccaaggt     780
gcagcattgc cactactgtg gcggctccgg catggaaacc atcaacacag gcccttttgt     840
gatgcgttcc acgtgtagga gatgtggtgg ccgcggctcc atcatcatat cgccctgtgt     900
ggtctgcagg ggagcaggac aagccaagca gaaaaagcga gtgatgatcc ctgtgcctgc     960
aggagtcgag gatggccaga ccgtgaggat gcctgtggga aaaagggaaa ttttcattac    1020
gttcagggtg cagaaaagcc ctgtgttccg gagggacggc gcagacatcc actccgacct    1080
ctttatttct atagctcagg ctcttcttgg gggaacagcc agagcccagg gcctgtacga    1140
gacgatcaac gtgacgatcc cccctgggac tcagacagac cagaagattc ggatgggtgg    1200
gaaaggcatc cccccggatta acagctacgg ctacggagac cactacatcc acatcaagat    1260
acgagttcca aagaggctaa cgagccggca gcagagcctg atcctgagct acgccgagga    1320
cgagacagat gtggagggga cggtgaacgg cgtcaccctc accagctctg gtggcagcac    1380
catggatagc tccgcaggaa gcaaggctag gcgtgaggct ggggaggacg aggagggatt    1440
cctttccaaa cttaagaaaa tgtttacctc atgatatccc agccgaggaa aaagatccac    1500
tggaaactag gccgggaagc agcagcccct ccaagggcca gggcacctgg gagacgggag    1560
gattccagaa cagcagcact gagctcccac ccgcagagcc tctggacggc cttggcaaca    1620
gcaaaatcat gggacaacac ctctctccac ggaaaggtca cagtggacag cccgggcagt    1680
aggatgcagc cccagaggct ggtggcagtt tcctgtccat tggtaggtga cggccccctg    1740
gtcagcagag gagaggttag atcttgcagg ctaaaactct aatttggaat tgaatattgt    1800
ggatatctta gttaaaggcc atgcttacag cttagaaatg aagccttaag ctgcatcaag    1860
ttacgaagtg attaatttcc ttctcagcaa acctccggga ggttccagaa tgagttcttc    1920
```

ctgacaggtt gtcttcactg ggagcgtggg gcccccaggc cccaccagca ccgtcctccc 1980
ctaatgaggg gccctgccga ggcatcagct gctctgctca gttagttttt attcccgggg 2040
taccaagcag ctgcacagtc ggtgcctggg aagcacgtta aaggcccaga gagatcctgg 2100
gggttctgct ctgaccgtgt gggtggtgat ccttgtcagg atgtacagtc cttgctccca 2160
ccccatccgg gatggccgcc tgtccctgac tattgagtcc tgttgttgta agccaggcat 2220
ggagggctcc tgcccttctg ctgagccaca gcccattgca gcactgtgct ggccagactt 2280
cagctgcctt gggaactgaa gccctgccac tgttgctagt caggggcttg gttctcccac 2340
ttacactgtt gacatctatt ttctgaagtg tgtttaaatt attcagtgct aatcattgtt 2400
ttttcctttg taaatgttga ttcagaaaag gaaagcacag gctaagcagt tgaaggttcc 2460
ccaccattca gtgagagcag aacccccatt ccccagcctc tgctggtagc atgtcgcagt 2520
ttccatgtgt ttcaggatct tcgggctgtc gttagacagg ttaatgaaga acacttctca 2580
acagtttcct ttttgttttc ctttataatt cactaaaata aagcatctat tagtgtctga 2640
aaaaaaaaaa aaaaaa 2656

<210> 84
<211> 2217
<212> DNA
<213> Homo sapiens

<400> 84
gcggacccgg cgccgaggcg gccacccgag acgcggcgcg cacgctccgg cctgcgcagc 60
ccggcccggc catggcggcc ccccgcccgt ctcccgcgat ctccgtttcg gtctcggctc 120
cggcttttta cgccccgcag aagaagttcg gccctgtggt ggccccaaag cccaaagtga 180
atcccttccg gcccggggac agcgagcctc ccccggcacc cggggcccag cgcgcacaga 240
tgggccgggt gggcgagatt cccccgccgc ccccggaaga ctttcccctg cctccacctc 300
cccttgctgg ggatggcgac gatgcagagg gtgctctggg aggtgccttc ccgccgcccc 360
ctcccccgat cgaggaatca tttcccccctg cgcctctgga ggaggagatc ttcccttccc 420
cgccgcctcc tccggaggag gagggagggc ctgaggcccc cataccgccc ccaccacagc 480
ccagggagaa ggtgagcagt attgatttgg agatcgactc tctgtcctca ctgctggatg 540
acatgaccaa gaatgatcct ttcaaagccc gggtgtcatc tggatatgtg cccccaccag 600
tggccactcc attcagttcc aagtccagta ccaagcctgc agccgggggc acagcacccc 660
tgcctccttg gaagtcccct tccagctccc agcctctgcc ccaggttccg gctccggctc 720
agagccagac acagttccat gttcagcccc agccccagcc caagcctcag gtccaactcc 780
atgtccagtc ccagacccag cctgtgtctt tggctaacac ccagccccga gggcccccag 840
cctcatctcc ggctccagcc cctaagtttt ctccagtgac tcctaagttt actcctgtgg 900
cttccaagtt cagtcctgga gccccaggtg gatctgggtc acaaccaaat caaaaattgg 960
ggcaccccga agctctttct gctggcacag gctcccctca acctcccagc ttcacctatg 1020
cccagcagag ggagaagccc cgagtgcagg agaagcagca ccccgtgccc ccaccggctc 1080
agaaccaaaa ccaggtgcgc tcccctgggg ccccagggcc ccctgactctg aaggaggtgg 1140
aggagctgga gcagctgacc cagcagctaa tgcaggacat ggagcatcct cagaggcaga 1200
atgtggctgt caacgaactc tgcggccgat gccatcaacc cctggcccgg gcgcagccag 1260
ccgtccgcgc tctagggcag ctgttccaca tcgcctgctt cacctgccac cagtgtgcgc 1320
agcagctcca gggccagcag ttctacagtc tggagggggc gccgtactgc gagggctgtt 1380
acactgacac cctggagaag tgtaacacct gcggggagcc catcactgac cgcatgctga 1440
gggccacggg caaggcctat cacccgcact gcttcacctg tgtggtctgc gcccgccccc 1500
tggagggcac ctccttcatc gtggaccagg ccaaccggcc ccactgtgtc cccgactacc 1560
acaagcagta cgcccccgagg tgctccgtct gctctgagcc catcatgcct gagcctggcc 1620
gagatgagac tgtgcgagtg gtcgccctgg acaagaactt ccacatgaag tgttacaagt 1680
gtgaggactg cgggaagccc ctgtcgattg aggcagatga caatggctgc ttcccccctgg 1740
acggtcacgt gctctgtcgg aagtgccaca ctgctagagc ccagacctga gtgaggacag 1800
gccctcttca gaccgcagtc catgccccat tgtggaccac ccacactgag accacctgcc 1860
cccacctcag ttattgtttt gatgtctagc ccctcccatt tccaacccct ccctagcatc 1920
ccaggtgccc tgacccagga cccaacatgg tctagggatg caggatcccc gccctggggt 1980
ctggtcctcg cccatcctgc agggattgcc caccgtcttc cagacacccc acctgagggg 2040
ggcaccaggt ttagtgctgc tgctttcact gctgcacccg cgccctcggc cggcccccccg 2100
agcagccttt gtactctgct tgcggagggc tgggagaccc tccaggacat tcccaccctc 2160
ccccatgctg ccaagttgta gctatagcta caaataaaaa aaaaccttgt tttccag 2217

<210> 85
<211> 8906
<212> DNA
<213> Homo sapiens

<400> 85
gaggcggcca aggacctggc cgacatcgcg gccttcttcc gatccgggtt tcgaaaaaac 60
gatgaaatga aagctatgga tgttttacca attttgaagg aaaaagttgc atacctttca 120
ggtgggagag ataaacgtgg aggtcccatt ttaacgtttc cggcccgcag caatcatgac 180
agaatacgac aggaggatct caggagactc atttcctatc tagcctgtat tcccagcgag 240
gaggtctgca agcgtggctt cacggtgatc gtggacatgc gtgggtccaa gtgggactcc 300
atcaagcccc ttctgaagat cctgcaggag tccttcccct gctgcatcca tgtggccctg 360
atcatcaagc cagacaactt ctggcagaaa cagaggacta attttggcag ttctaaattt 420
gaatttgaga caaatatggt ctctttagaa ggccttacca aagtagttga tccttctcag 480
ctaactcctg agtttgatgg ctgcctggaa tacaaccacg aagaatggat tgaaatcaga 540
gttgcttttg aagactacat tagcaatgcc acccacatgc tgtctcggct ggaggaactt 600
caggacatcc tagctaagaa ggagctgcct caggatttag aggggggctcg gaatatgatc 660
gaggaacatt ctcagctgaa gaagaaggtg attaaggccc ccatcgagga cctggatttg 720
gagggacaga agctgcttca gaggatacag agcagtgaaa gctttcccaa aaagaactca 780
ggctcaggca atgcggacct gcagaacctc ttgcccaagg tgtccaccat gctggaccgg 840
ctgcactcga cacggcagca tctgcaccag atgtggcatg tgaggaagct gaagctggac 900
cagtgcttcc agctgaggct gtttgaacag gatgctgaga agatgtttga ctggatcaca 960
cacaacaaag gcctgtttct aaacagctac acagagattg ggaccagcca ccctcatgcc 1020
atggagcttc agacgcagca caatcacttt gccatgaact gtatgaacgt gtatgtaaat 1080
ataaaccgca tcatgtcggt ggccaatcgt ctggtggagt ctggccacta tgcctcgcag 1140
cagatcaggc agatcgcgag tcagctggag caggagtgga aggcgtttgc ggcagccctg 1200
gatgagcgga gcaccttgct ggacatgtcc tccattttcc accagaaggc cgaaaagtat 1260
atgagcaacg tggattcatg gtgtaaagct tgcggtgagg tagaccttcc ctcagagctg 1320
caggacctag aagatgccat tcatcaccac cagggaatat atgaacatat cactcttgct 1380
tattctgagg tcagccaaga tgggaagtcg ctccttgaca agctccagcg gcccttgact 1440
cccggcagct ccgattccct gacagcctct gccaactact ccaaggccgt gcaccatgtc 1500
ctggatgtca tccacgaggt gctgcaccac cagcggcacg tgagaacaat ctggcaacac 1560
cgcaaggtcc ggctgcatca gaggctgcag ctgtgtgttt tccagcagga agttcagcag 1620
gtgctagact ggatcgagaa ccacggagaa gcatttctga gcaaacatac aggtgtgggg 1680
aaatctcttc atcgggccag agcattgcag aaacgtcatg aagattttga agaagtggca 1740
cagaacacat acaccaatgc ggataaatta ctggaagcag cagaacagct ggctcagact 1800
ggggaatgtg accccgaaga gatttatcag gctgcccatc agctggaaga ccggattcaa 1860
gatttcgttc ggcgtgttga gcagcgaaag atcctactgg acatgtcagt gtcctttcac 1920
acccatgtga aagagctgtg gacgtggctg gaggagctgc agaaggagct gctggacgac 1980
gtgtatgccg agtcggtgga ggccgtgcag gacctcatca agcgctttgg ccagcagcag 2040
cagaccaccc tgcaggtgac tgtcaacgtg atcaaggaag gggaggacct catccagcag 2100
ctcagggact ctgccatctc cagtaacaag accccccaca acagctccat caaccacatt 2160
gagacggtgc tgcagcagct ggacgaggcg cagtcgcaga tggaggagct cttccaggag 2220
cgcaagatca agctggagct cttcctgcac gtgcgcatct tcgagaggga cgccatcgac 2280
attatctcag acctcgagtc ttggaatgat gagctttctc agcaaatgaa tgacttcgac 2340
acagaagatc tcacgattgc agagcagcgc ctccagcacc atgcagacaa agccttgacc 2400
atgaacaact tgacttttga cgtcatccac caagggcaag atcttctgca gtatgtcaat 2460
gaggtccagg cctctggtgt ggagctgctg tgtgatagag atgtagacat ggcaactcgg 2520
gtccaggacc tgctggagtt tcttcatgaa aaacagcagg aattggattt agccgcagag 2580
cagcatcgga aacacctgga gcagtgcgtg cagctgcgcc acctgcaggc agaagtgaaa 2640
caggtgctgg gttggatccg caacggagag tccatgttaa atgccggact tatcacagcc 2700
agctcgttac aagaggcaga gcagctccag cgagagcacg agcagttcca gcatgccatt 2760
gagaaaacac atcagagcgc gctgcaggtg cagcagaagg cagaagccat gctacaggcc 2820
aaccactacg acatggacat gatccgggac tgcgccgaga aggtggcgtc tcactggcaa 2880
cagctcatgc tcaagatgga agatcgcctc aagctcgtca acgcctctgt cgctttctac 2940
aaaacctcag agcaggtctg cagcgtcctc gagagcctgg aacaggagta caagagagaa 3000

```
gaagactggt gtggcggggc ggataagctg ggcccaaact ctgagacgga ccacgtgacg    3060
cccatgatca gcaagcacct ggagcagaag gaggcattcc tgaaggcttg caccccttgct   3120
cggaggaatg cagacgtctt cctgaaatac ctgcacagga acagcgtgaa catgccagga    3180
atggtgacgc acatcaaagc tcctgaacag caagtgaaaa atatcttgaa tgaactcttc    3240
caacgggaga acagggtatt gcattactgg accatgagga agagacggct ggaccagtgt    3300
cagcagtacg tggtctttga gaggagtgcc aagcaggctt tggaatggat ccatgacaat    3360
ggcgagttct acctttccac acacacctcc acgggctcca gtatacagca cacccaggag    3420
ctcctgaaag agcacgagga gttccagata actgcaaagc aaaccaaaga gagagtgaag    3480
ctattgatac agctggctga tggcttttgt gaaaaagggc atgcccatgc ggcagagata    3540
aaaaaatgtg ttactgctgt ggataagagg tacagagatt tctctctgcg gatggagaag    3600
tacaggacct ctttggaaaa agccctgggg atttcttcag attccaacaa atcgagtaaa    3660
agtctccagc tagatatcat tccagccagt atccctggct cagaggtgaa acttcgagat    3720
gctgctcatg aacttaatga agagaagcgg aaatctgccc gcaggaaaga gttcataatg    3780
gctgagctca ttcaaactga aaaggcttat gtaagagacc tccgggaatg tatggatacg    3840
tacctgtggg aaatgaccag tggcgtggaa gagattccac ctggcattgt aaacaaagaa    3900
ctcatcatct tcggaaacat gcaagaaatc tacgaatttc ataataacat attcctaaag    3960
gagctggaaa aatatgaaca gttgccagag gatgttggac attgttttgt tacttgggca    4020
gacaagtttc agatgtatgt cacatattgc aaaaataagc ctgattctac tcagctgata    4080
ttggaacatg cagggtccta ttttgacgag atacagcagc gacatggatt agccaattcc    4140
atttcttcct accttattaa accagttcag cgaataacga aatatcagct ccttttaaaa    4200
gagctgctga cgtgctgtga ggaaggaaag ggagagatta aagatggcct ggaggtgatg    4260
ctcagcgtgc cgaagcgagc caatgacgcc atgcacctca gcatgctgga agggtttgat    4320
gaaaacattg agtctcaggg agaactcatc ctacaggaat ccttccaagt gtgggaccca    4380
aaaaccttaa ttcgaaaggg tcgagaacgg catctcttcc tttttgaaat gtccttagta    4440
tttagtaaag aagtgaaaga ttccagtggg agaagcaagt acctttataa aagcaaattg    4500
tttacctcag agttgggtgt cacagaacat gttgaaggag acccttgcaa atttgcactg    4560
tgggtgggga gaacaccaac ttcagataat aaaattgtcc ttaaggcttc cagcatagag    4620
aacaagcagg actggataaa gcatatccgc gaagtcatcc aggagcggac gatccacctg    4680
aagggagccc tgaaggagcc cattcacatc cctaagaccg ctcccgccac aagacagaag    4740
ggaaggaggg atggagagga tctggacagc caaggagacg gcagcagcca gcctgatacg    4800
atttccatcg cctcacggac gtctcagaac acgctggaca gcgataagct ctctggtggc    4860
tgtgagctga cagtggtgat ccatgacttc accgcttgca acagcaacga gctgaccatc    4920
cgacggggcc agaccgtgga agttctggag cggccgcatg acaagcctga ctggtgtctg    4980
gtgcggacca ctgaccgctc cccagcggca gaaggcctgg tcccctgtgg ttcactgtgc    5040
atcgcccact ccagaagtag catggaaatg gagggcatct tcaaccacaa agactcgctc    5100
tccgtctcca gcaatgacgc cagtccaccc gcatccgtgg cttccctcca gccccacatg    5160
atcgggggccc agagctcgcc gggccccaag cggccgggca acaccctgcg caagtggctc   5220
accagccccg tgcggcggct cagcagcggc aaggccgacg ggcacgtgaa gaagctggcg    5280
cacaagcaca agaagagccg cgaggtccgc aagagcgccg acgccggctc gcagaaggac    5340
tccgacgaca gtgcggccac cccgcaggac gagacggtcg aggagagagg ccggaacgag    5400
ggcctgagca gcggtactct ctccaaatcc tcctcctcgg ggatgcagag ctgtggagaa    5460
gaggaaggcg aggaggggggc cgacgccgtg cccctgccgc cacccatggc catccagcag   5520
cacagcctcc tccagccaga ctcacaggat gacaaggcct cttctcggtt attagtccgc    5580
cccaccagct ccgaaacacc gagtgcagcc gagctcgtca gtgcaattga ggaactcgtg    5640
aaaagcaaga tggcactgga ggatcgcccc agctcactcc ttgttgacca gggagatagt    5700
agcagccctt ccttcaaccc ttcggataat tcccttctct cttcctcctc gcccattgat    5760
gagatggaag aaaggaaatc cagctcttta aagagaagac actacgtttt gcaagaacta    5820
gtggagacag agcgtgacta tgtgcgggac cttggctatg tggttgaggg ctacatggca    5880
cttatgaaag aagatggtgt tcctgatgac atgaaaggaa aagacaaaat tgtgttcggc    5940
aacatccatc agatttacga ctggcacaga gacttttttt taggagagtt agagaagtgc    6000
cttgaagatc cagaaaaact aggatccctt tttgttaaac acgagagaag gttgcacatg    6060
tacatagctt attgtcaaaa taaaccaaag tctgagcaca ttgtctcaga atacattgat    6120
accttttttg aggacttaaa gcagcgtctt ggccacaggt tacagctcac agatctgttg    6180
atcaaaccag tgcagagaat catgaagtat cagctgttac tgaaggactt cctcaagtat    6240
tccaaaaagg ccagcctgga tacatcagaa ttagagagag ctgtggaagt catgtgcata    6300
gtacccaggc ggtgcaacga catgatgaac gtggggcggc tgcaaggatt cgacgggaaa    6360
atcgttgccc agggtaaact gctcttgcag gacacattct tggtcacaga ccaagatgca    6420
```

```
ggacttctgc ctcgctgcag agagaggcgc atcttcctct ttgagcagat cgtcatattc    6480
agcgaaccac ttgataaaaa gaagggcttc tccatgccgg gattcctgtt taagaacagt    6540
atcaaggtga gttgcctttg cctggaggaa aatgtggaaa atgatccctg taaatttgct    6600
ctgacatcga ggacgggtga cgtggtagag accttcattt tgcattcatc tagtccaagt    6660
gtccggcaaa cttggatcca tgaaatcaac caaattttag aaaaccagcg caattttta    6720
aatgccttga catcgccaat cgagtaccag aggaaccaca gcgggggcgg cggcggcggc    6780
ggcagcgggg cagcggcggg ggtgggggca gcggcggcgg cggggccccc agtggcggca    6840
gcggccacag tggcggcccc agcagctgcg gcggcgcccc cagcacgagc aggagccggc    6900
cctcccggat cccccagcct gtccgacacc acccccccgt gctggtctcc tctgcagcct    6960
cgagccaggc agaggcagac aagatgtcag agtgaaagca gcagcagtag caacatctcc    7020
accatgttgg tgacacacga ttacacggca gtgaaggagg atgagatcaa cgtctaccaa    7080
ggagaggtcg ttcaaattct ggccagcaac cagcagaaca tgtttctggt gttccgagcc    7140
gccactgacc agtgccccgc agctgagggc tggattccag gctttgtcct gggccacacc    7200
agtgcagtca tcgtggagaa cccggacggg actctcaaga agtcaacatc ttggcacaca    7260
gcactccgtt taaggaaaaa atctgagaaa aaagataaag acggcaaaag ggaaggcaag    7320
ttagagaacg gttatcggaa gtcacgggaa ggactcagca acaaggtatc tgtgaagctt    7380
ctcaatccca actacattta tgacgttccc ccagaattcg tcattccatt gagtgaggtc    7440
acgtgtgaga caggggagac cgttgttctt agatgtcgag tctgtggccg ccccaaagcc    7500
tcaattacct ggaagggccc tgaacacaac accttgaaca acgatggtca ctacagcatc    7560
tcctacagtg acctgggaga ggccacgctg aagattgtgg gcgtgaccac ggaagatgac    7620
ggcatctaca cgtgcatcgc tgtcaatgac atgggttcag cctcatcatc ggccagcctg    7680
agggtcctag gtccagggat ggatgggatc atggtgacct ggaaagacaa ctttgactcc    7740
ttctacagtg aagtggctga gcttggcagg ggcagattct ctgtcgttaa gaaatgtgat    7800
cagaaaggaa ccaagcgagc agtggccact aagtttgtga acaagaagtt gatgaagcgc    7860
gaccaggtca cccatgagct tggcatcctg cagagcctcc agcacccccct gcttgtcggc    7920
ctcctcgaca cctttgagac cccccaccagc tacatcctgg tcttagaaat ggctgaccag    7980
ggtcgcctcc tggactgcgt ggtgcgatgg ggaagcctca ctgaagggaa gatcagggcg    8040
cacctggggg aggttctgga agctgtccgg tacctgcaca actgcaggat agcacacctg    8100
gacctaaagc ctgagaatat cctggtggat gagagtttag ccaagccaac catcaaactg    8160
gctgactttg gagatgctgt tcagctcaac acgacctact acatccacca gttactgggg    8220
aaccctgaat tcgcagcccc tgaaatcatc ctcgggaacc ctgtctccct gacctcggat    8280
acgtggagtg ttggagtgct cacatacgta cttcttagtg gcgtgtcccc cttcctggat    8340
gacagtgtgg aagagacctg cctgaacatt tgccgcttag actttagctt cccagatgac    8400
tactttaaag gagtgagcca gaaggccaag gagttcgtgt gcttcctcct gcaggaggac    8460
cccgccaagc gtccctcggc tgcgctggcc ctccaggagc agtggctgca ggccggcaac    8520
ggcagaagca cgggcgtcct cgacacgtcc agactgactt ccttcattga gcggcgcaaa    8580
caccagaatg atgttcgacc tatccgtagc attaaaaact ttctgcagag caggcttctg    8640
cctagagttt gacctatcca gaagttcttt ctcattctct ttcacctgcc aatcagctgt    8700
taatctgaat tttcaagaga aaacaagcaa acataactga tcagctgccg gtatgttcat    8760
cgtgtgaaat tgcattccaa gtgagctgtg ctcagcagtg cttggacaca gagctgcaag    8820
ctgcgctggg gtggaggacc gtcacttaca ctctgccaag gacggaggtc gcattgctgt    8880
atcacagtat tttttacgga tttctg                                          8906
```

<210> 86
<211> 1204
<212> DNA
<213> Homo sapiens

<400> 86

```
tcggcggcgg tggtatcggc ggcagctgtg aggggggttcc gggaagatgg tgctgatcaa      60
ggaattccgt gtggttttgc catgttctgt tcaggagtat caggttgggc agctttactc     120
tgttgcagaa gctagtaaga atgagactgg tggtggagaa ggaattgaag tcttaaagaa     180
tgaaccttat gagaaggatg gagaaaaggg acagtatacg cacaaaattt atcacctaaa     240
gagcaaagtg cctgcattcg tgaggatgat tgctcccgag ggctccttgg tgtttcatga     300
gaaagcctgg aatgcgtacc cctactgtag aacaattgta acgaatgaat atatgaaaga     360
tgatttcttc attaaaatcg aaacatggca caaaccagac ttgggaacat tagaaaatgt     420
```

acatggttta gatccaaaca catggaaaac tgttgaaatt gtccatatag atattgcaga 480
tagaagtcaa gttgaaccag cagactacaa agctgatgaa gacccagcat tattccagtc 540
agtcaagacc aagagaggcc ctttgggacc caactggaag aaggagctgg caaacagccc 600
tgactgtccc cagatgtgtg cctataagct ggtgaccatc aaattcaagt ggtggggact 660
gcaaagcaaa gtagaaaact tcattcaaaa gcaagaaaaa cggatattta caaacttcca 720
tcgccagctt ttttgttgga ttgacaagtg gatcgatctc acgatggaag acattaggag 780
aatggaagac gagactcaga aagaactaga aacaatgcgt aagaggggtt ccgttcgagg 840
cacgtcggct gctgatgtct agatgagtcc cctgtagggt cagagacaat gtcaaactgt 900
ttacgtaatc aaggtcaagt gaggggaaca agcgcagcca gtgatgagtg aacaacaatc 960
tgaccagtat cttgcagtgt tgacgtttcc cagatgtgtg cttgtgatga tacacacaca 1020
tgcacaggtt ctcaaccacg tgtgtatata tgtatgtgtg catatgtctg tagctgtata 1080
taaagcgcat gtagagctac agatccagat acacacactt gtgtatatat gtacatacag 1140
acatactgaa gggattagta caatttctcc aaagtactgt acctatcttc agcaagaatg 1200
caaa 1204

<210> 87
<211> 892
<212> PRT
<213> Homo sapiens

<400> 87
Met Asp His Tyr Asp Ser Gln Gln Thr Asn Asp Tyr Met Gln Pro Glu
1 5 10 15
Glu Asp Trp Asp Arg Asp Leu Leu Leu Asp Pro Ala Trp Glu Lys Gln
20 25 30
Gln Arg Lys Thr Phe Thr Ala Trp Cys Asn Ser His Leu Arg Lys Ala
35 40 45
Gly Thr Gln Ile Glu Asn Ile Glu Glu Asp Phe Arg Asp Gly Leu Lys
50 55 60
Leu Met Leu Leu Leu Glu Val Ile Ser Gly Glu Arg Leu Ala Lys Pro
65 70 75 80
Glu Arg Gly Lys Met Arg Val His Lys Ile Ser Asn Val Asn Lys Ala
85 90 95
Leu Asp Phe Ile Ala Ser Lys Gly Val Lys Leu Val Ser Ile Gly Ala
100 105 110
Glu Glu Ile Val Asp Gly Asn Val Lys Met Thr Leu Gly Met Ile Trp
115 120 125
Thr Ile Ile Leu Arg Phe Ala Ile Gln Asp Ile Ser Val Glu Glu Thr
130 135 140
Ser Ala Lys Glu Gly Leu Leu Leu Trp Cys Gln Arg Lys Thr Ala Pro
145 150 155 160
Tyr Lys Asn Val Asn Ile Gln Asn Phe His Ile Ser Trp Lys Asp Gly
165 170 175
Leu Gly Phe Cys Ala Leu Ile His Arg His Arg Pro Glu Leu Ile Asp
180 185 190
Tyr Gly Lys Leu Arg Lys Asp Asp Pro Leu Thr Asn Leu Asn Thr Ala
195 200 205
Phe Asp Val Ala Glu Lys Tyr Leu Asp Ile Pro Lys Met Leu Asp Ala
210 215 220
Glu Asp Ile Val Gly Thr Ala Arg Pro Asp Glu Lys Ala Ile Met Thr
225 230 235 240
Tyr Val Ser Ser Phe Tyr His Ala Phe Ser Gly Ala Gln Lys Ala Glu
245 250 255
Thr Ala Ala Asn Arg Ile Cys Lys Val Leu Ala Val Asn Gln Glu Asn
260 265 270
Glu Gln Leu Met Glu Asp Tyr Glu Lys Leu Ala Ser Asp Leu Leu Glu
275 280 285

Trp Ile Arg Arg Thr Ile Pro Trp Leu Glu Asn Arg Val Pro Glu Asn
290 295 300
Thr Met His Ala Met Gln Gln Lys Leu Glu Asp Phe Arg Asp Tyr Arg
305 310 315 320
Arg Leu His Lys Pro Pro Lys Val Gln Glu Lys Cys Gln Leu Glu Ile
325 330 335
Asn Phe Asn Thr Leu Gln Thr Lys Leu Arg Leu Ser Asn Arg Pro Ala
340 345 350
Phe Met Pro Ser Glu Gly Arg Met Val Ser Asp Ile Asn Asn Ala Trp
355 360 365
Gly Cys Leu Glu Gln Val Glu Lys Gly Tyr Glu Glu Trp Leu Leu Asn
370 375 380
Glu Ile Arg Arg Leu Glu Arg Leu Asp His Leu Ala Glu Lys Phe Arg
385 390 395 400
Gln Lys Ala Ser Ile His Glu Ala Trp Thr Asp Gly Lys Glu Ala Met
405 410 415
Leu Arg Gln Lys Asp Tyr Glu Thr Ala Thr Leu Ser Glu Ile Lys Ala
420 425 430
Leu Leu Lys Lys His Glu Ala Phe Glu Ser Asp Leu Ala Ala His Gln
435 440 445
Asp Arg Val Glu Gln Ile Ala Ala Ile Ala Gln Glu Leu Asn Glu Leu
450 455 460
Asp Tyr Tyr Asp Ser Pro Ser Val Asn Ala Arg Cys Gln Lys Ile Cys
465 470 475 480
Asp Gln Trp Asp Asn Leu Gly Ala Leu Thr Gln Lys Arg Arg Glu Ala
485 490 495
Leu Glu Arg Thr Glu Lys Leu Leu Glu Thr Ile Asp Gln Leu Tyr Leu
500 505 510
Glu Tyr Ala Lys Arg Ala Ala Pro Phe Asn Asn Trp Met Glu Gly Ala
515 520 525
Met Glu Asp Leu Gln Asp Thr Phe Ile Val His Thr Ile Glu Glu Ile
530 535 540
Gln Gly Leu Thr Thr Ala His Glu Gln Phe Lys Ala Thr Leu Pro Asp
545 550 555 560
Ala Asp Lys Glu Arg Leu Ala Ile Leu Gly Ile His Asn Glu Val Ser
565 570 575
Lys Ile Val Gln Thr Tyr His Val Asn Met Ala Gly Thr Asn Pro Tyr
580 585 590
Thr Thr Ile Thr Pro Gln Glu Ile Asn Gly Lys Trp Asp His Val Arg
595 600 605
Gln Leu Val Pro Arg Arg Asp Gln Ala Leu Thr Glu Glu His Ala Arg
610 615 620
Gln Gln His Asn Glu Arg Leu Arg Lys Gln Phe Gly Ala Gln Ala Asn
625 630 635 640
Val Ile Gly Pro Trp Ile Gln Thr Lys Met Glu Glu Ile Gly Arg Ile
645 650 655
Ser Ile Glu Met His Gly Thr Leu Glu Asp Gln Leu Ser His Leu Arg
660 665 670
Gln Tyr Glu Lys Ser Ile Val Asn Tyr Lys Pro Lys Ile Asp Gln Leu
675 680 685
Glu Gly Asp His Gln Leu Ile Gln Glu Ala Leu Ile Phe Asp Asn Lys
690 695 700
His Thr Asn Tyr Thr Met Glu His Ile Arg Val Gly Trp Glu Gln Leu
705 710 715 720
Leu Thr Thr Ile Ala Arg Thr Ile Asn Glu Val Glu Asn Gln Ile Leu
725 730 735
Thr Arg Asp Ala Lys Gly Ile Ser Gln Glu Gln Met Asn Glu Phe Arg

```
            740                 745                 750
Ala Ser Phe Asn His Phe Asp Arg Asp His Ser Gly Thr Leu Gly Pro
        755                 760                 765
Glu Glu Phe Lys Ala Cys Leu Ile Ser Leu Gly Tyr Asp Ile Gly Asn
    770                 775                 780
Asp Pro Gln Gly Glu Ala Glu Phe Ala Arg Ile Met Ser Ile Val Asp
785                 790                 795                 800
Pro Asn Arg Leu Gly Val Val Thr Phe Gln Ala Phe Ile Asp Phe Met
                805                 810                 815
Ser Arg Glu Thr Ala Asp Thr Asp Thr Ala Asp Gln Val Met Ala Ser
            820                 825                 830
Phe Lys Ile Leu Ala Gly Asp Lys Asn Tyr Ile Thr Met Asp Glu Leu
        835                 840                 845
Arg Arg Glu Leu Pro Pro Asp Gln Ala Glu Tyr Cys Ile Ala Arg Met
    850                 855                 860
Ala Pro Tyr Thr Gly Pro Asp Ser Val Pro Gly Ala Leu Asp Tyr Met
865                 870                 875                 880
Ser Phe Ser Thr Ala Leu Tyr Gly Glu Ser Asp Leu
                885                 890
```

<210> 88
<211> 197
<212> PRT
<213> Homo sapiens

<400> 88

```
Met Met Phe Pro Gln Ser Arg His Ser Gly Ser Ser His Leu Pro Gln
1               5                   10                  15
Gln Leu Lys Phe Thr Thr Ser Asp Ser Cys Asp Arg Ile Lys Asp Glu
            20                  25                  30
Phe Gln Leu Leu Gln Ala Gln Tyr His Ser Leu Lys Leu Glu Cys Asp
        35                  40                  45
Lys Leu Ala Ser Glu Lys Ser Glu Met Gln Arg His Tyr Val Met Tyr
    50                  55                  60
Tyr Glu Met Ser Tyr Gly Leu Asn Ile Glu Met His Lys Gln Ala Glu
65                  70                  75                  80
Ile Val Lys Arg Leu Asn Gly Ile Cys Ala Gln Val Leu Pro Tyr Leu
                85                  90                  95
Ser Gln Glu His Gln Gln Gln Val Leu Gly Ala Ile Glu Arg Ala Lys
            100                 105                 110
Gln Val Thr Ala Pro Glu Leu Asn Ser Ile Ile Arg Gln Gln Leu Gln
        115                 120                 125
Ala His Gln Leu Ser Gln Leu Gln Ala Leu Ala Leu Pro Leu Thr Pro
    130                 135                 140
Leu Pro Val Gly Leu Gln Pro Pro Ser Leu Pro Ala Val Ser Ala Gly
145                 150                 155                 160
Thr Gly Leu Leu Ser Leu Ser Ala Leu Gly Ser Gln Ala His Leu Ser
                165                 170                 175
Lys Glu Asp Lys Asn Gly His Asp Gly Asp Thr His Gln Glu Asp Asp
            180                 185                 190
Gly Glu Lys Ser Asp
        195
```

<210> 89
<211> 739
<212> PRT
<213> Homo sapiens

<400> 89
Gly Asp Lys Glu Pro Thr Glu Thr Ile Gly Asp Leu Ser Ile Cys Leu
1   5   10   15
Asp Gly Leu Gln Leu Glu Ser Glu Val Val Thr Asn Gly Glu Thr Thr
            20   25   30
Cys Ser Glu Ser Ala Ser Gln Asn Asp Asp Gly Ser Arg Ser Lys Asp
        35   40   45
Glu Thr Arg Val Ser Thr Asn Gly Ser Asp Asp Pro Glu Asp Ala Gly
    50   55   60
Ala Gly Glu Asn Arg Arg Val Ser Gly Asn Asn Ser Pro Ser Leu Ser
65   70   75   80
Asn Gly Gly Phe Lys Pro Ser Arg Pro Pro Arg Pro Ser Arg Pro Pro
                85   90   95
Pro Pro Thr Pro Arg Arg Pro Ala Ser Val Asn Gly Ser Pro Ser Ala
            100   105   110
Thr Ser Glu Ser Asp Gly Ser Ser Thr Gly Ser Leu Pro Pro Thr Asn
        115   120   125
Thr Asn Thr Asn Thr Ser Glu Gly Ala Thr Ser Gly Leu Ile Ile Pro
    130   135   140
Leu Thr Ile Ser Gly Gly Ser Gly Pro Arg Pro Leu Asn Pro Val Thr
145   150   155   160
Gln Ala Pro Leu Pro Pro Gly Trp Glu Gln Arg Val Asp Gln His Gly
                165   170   175
Arg Val Tyr Tyr Val Asp His Val Glu Lys Arg Thr Thr Trp Asp Arg
            180   185   190
Pro Glu Pro Leu Pro Pro Gly Trp Glu Arg Arg Val Asp Asn Met Gly
        195   200   205
Arg Ile Tyr Tyr Val Asp His Phe Thr Arg Thr Thr Thr Trp Gln Arg
    210   215   220
Pro Thr Leu Glu Ser Val Arg Asn Tyr Glu Gln Trp Gln Leu Gln Arg
225   230   235   240
Ser Gln Leu Gln Gly Ala Met Gln Gln Phe Asn Gln Arg Phe Ile Tyr
                245   250   255
Gly Asn Gln Asp Leu Phe Ala Thr Ser Gln Ser Lys Glu Phe Asp Pro
            260   265   270
Leu Gly Pro Leu Pro Pro Gly Trp Glu Lys Arg Thr Asp Ser Asn Gly
        275   280   285
Arg Val Tyr Phe Val Asn His Asn Thr Arg Ile Thr Gln Trp Glu Asp
    290   295   300
Pro Arg Ser Gln Gly Gln Leu Asn Glu Lys Pro Leu Pro Glu Gly Trp
305   310   315   320
Glu Met Arg Phe Thr Val Asp Gly Ile Pro Tyr Phe Val Asp His Asn
                325   330   335
Arg Arg Thr Thr Thr Tyr Ile Asp Pro Arg Thr Gly Lys Ser Ala Leu
            340   345   350
Asp Asn Gly Pro Gln Ile Ala Tyr Val Arg Asp Phe Lys Ala Lys Val
        355   360   365
Gln Tyr Phe Arg Phe Trp Cys Gln Gln Leu Ala Met Pro Gln His Ile
    370   375   380
Lys Ile Thr Val Thr Arg Lys Thr Leu Phe Glu Asp Ser Phe Gln Gln
385   390   395   400
Ile Met Ser Phe Ser Pro Gln Asp Leu Arg Arg Arg Leu Trp Val Ile
                405   410   415
Phe Pro Gly Glu Glu Gly Leu Asp Tyr Gly Gly Val Ala Arg Glu Trp
            420   425   430
Phe Phe Leu Leu Ser His Glu Val Leu Asn Pro Met Tyr Cys Leu Phe

```
        435                     440                     445
Glu Tyr Ala Gly Lys Asp Asn Tyr Cys Leu Gln Ile Asn Pro Ala Ser
    450                     455                     460
Tyr Ile Asn Pro Asp His Leu Lys Tyr Phe Arg Phe Ile Gly Arg Phe
465                     470                     475                 480
Ile Ala Met Ala Leu Phe His Gly Lys Phe Ile Asp Thr Gly Phe Ser
                485                     490                     495
Leu Pro Phe Tyr Lys Arg Ile Leu Asn Lys Pro Val Gly Leu Lys Asp
            500                     505                     510
Leu Glu Ser Ile Asp Pro Glu Phe Tyr Asn Ser Leu Ile Trp Val Lys
        515                     520                     525
Glu Asn Asn Ile Glu Glu Cys Asp Leu Glu Met Tyr Phe Ser Val Asp
    530                     535                     540
Lys Glu Ile Leu Gly Glu Ile Lys Ser His Asp Leu Lys Pro Asn Gly
545                     550                     555                 560
Gly Asn Ile Leu Val Thr Glu Glu Asn Lys Glu Glu Tyr Ile Arg Met
                565                     570                     575
Val Ala Glu Trp Arg Leu Ser Arg Gly Val Glu Glu Gln Thr Gln Ala
            580                     585                     590
Phe Phe Glu Gly Phe Asn Glu Ile Leu Pro Gln Gln Tyr Leu Gln Tyr
        595                     600                     605
Phe Asp Ala Lys Glu Leu Glu Val Leu Leu Cys Gly Met Gln Glu Ile
    610                     615                     620
Asp Leu Asn Asp Trp Gln Arg His Ala Ile Tyr Arg His Tyr Ala Arg
625                     630                     635                 640
Thr Ser Lys Gln Ile Met Trp Phe Trp Gln Phe Val Lys Glu Ile Asp
                645                     650                     655
Asn Glu Lys Arg Met Arg Leu Leu Gln Phe Val Thr Gly Thr Cys Arg
            660                     665                     670
Leu Pro Val Gly Gly Phe Ala Asp Leu Met Gly Ser Asn Gly Pro Gln
        675                     680                     685
Lys Phe Cys Ile Glu Lys Val Gly Lys Glu Asn Trp Leu Pro Arg Ser
    690                     695                     700
His Thr Cys Phe Asn Arg Leu Asp Leu Pro Pro Tyr Lys Ser Tyr Glu
705                     710                     715                 720
Gln Leu Lys Glu Lys Leu Leu Phe Ala Ile Glu Glu Thr Glu Gly Phe
                725                     730                     735
Gly Gln Glu


<210> 90
<211> 431
<212> PRT
<213> Homo sapiens

<400> 90
Gly Pro Pro Pro Thr Arg Ala Leu Pro Leu Pro Gln Ser Leu Pro Pro
1               5                   10                  15
Asp Phe Arg Leu Glu Pro Thr Ala Pro Ala Leu Ser Pro Arg Ser Ser
            20                  25                  30
Phe Ala Ser Ser Ser Ala Ser Asp Ala Ser Lys Pro Ser Ser Pro Arg
        35                  40                  45
Gly Ser Leu Leu Leu Asp Gly Ala Gly Ala Gly Gly Ala Gly Gly Ser
    50                  55                  60
Arg Pro Cys Ser Asn Arg Thr Ser Gly Ile Ser Met Gly Tyr Asp Gln
65                  70                  75                  80
Arg His Gly Ser Pro Leu Pro Ala Gly Pro Cys Leu Phe Gly Pro Pro
```

```
                85                  90                  95
Leu Ala Gly Ala Pro Ala Gly Tyr Ser Pro Gly Gly Val Pro Ser Ala
            100                 105                 110
Tyr Pro Glu Leu His Ala Ala Leu Asp Arg Leu Tyr Ala Gln Arg Pro
        115                 120                 125
Ala Gly Phe Gly Cys Gln Glu Ser Arg His Ser Tyr Pro Pro Ala Leu
    130                 135                 140
Gly Ser Pro Gly Ala Leu Ala Gly Ala Arg Val Gly Ala Ala Gly Pro
145                 150                 155                 160
Leu Glu Arg Arg Gly Ala Gln Pro Gly Arg His Ser Val Thr Gly Tyr
                165                 170                 175
Gly Asp Cys Ala Val Gly Ala Arg Tyr Gln Asp Glu Leu Thr Ala Leu
            180                 185                 190
Leu Arg Leu Thr Val Gly Thr Gly Gly Arg Glu Ala Gly Ala Arg Gly
        195                 200                 205
Glu Pro Ser Gly Ile Glu Pro Ser Gly Leu Glu Glu Pro Pro Gly Pro
    210                 215                 220
Phe Val Pro Glu Ala Ala Arg Ala Arg Met Arg Glu Pro Glu Ala Arg
225                 230                 235                 240
Glu Asp Tyr Phe Gly Thr Cys Ile Lys Cys Asn Lys Gly Ile Tyr Gly
                245                 250                 255
Gln Ser Asn Ala Cys Gln Ala Leu Asp Ser Leu Tyr His Thr Gln Cys
            260                 265                 270
Phe Val Cys Cys Ser Cys Gly Arg Thr Leu Arg Cys Lys Ala Phe Tyr
        275                 280                 285
Ser Val Asn Gly Ser Val Tyr Cys Glu Glu Asp Tyr Leu Phe Ser Gly
    290                 295                 300
Phe Gln Glu Ala Ala Glu Lys Cys Cys Val Cys Gly His Leu Ile Leu
305                 310                 315                 320
Glu Lys Ile Leu Gln Ala Met Gly Lys Ser Tyr His Pro Gly Cys Phe
                325                 330                 335
Arg Cys Ile Val Cys Asn Lys Cys Leu Asp Gly Ile Pro Phe Thr Val
            340                 345                 350
Asp Phe Ser Asn Gln Val Tyr Cys Val Thr Asp Tyr His Lys Asn Tyr
        355                 360                 365
Ala Pro Lys Cys Ala Ala Cys Gly Gln Pro Ile Leu Pro Ser Glu Gly
    370                 375                 380
Cys Glu Asp Ile Val Arg Val Ile Ser Met Asp Arg Asp Tyr His Phe
385                 390                 395                 400
Glu Cys Tyr His Cys Glu Asp Cys Arg Met Gln Leu Ser Asp Glu Glu
                405                 410                 415
Gly Cys Cys Cys Phe Pro Leu Asp Gly His Leu Leu Cys His Gly
            420                 425                 430
```

<210> 91
<211> 900
<212> PRT
<213> Homo sapiens

<400> 91

```
Gly Pro Gly Ser Arg His His Arg Ala Arg Asp Arg Leu Ile His Phe
1               5                   10                  15
Gly Ala Val Ser Thr Asp Val Leu Gly Cys Ser Ala His Cys Ser Leu
            20                  25                  30
Thr Gln Ser Pro Lys Met Asn Ile Gln Glu Gln Gly Phe Pro Leu Asp
        35                  40                  45
Leu Gly Ala Ser Phe Thr Glu Asp Ala Pro Arg Pro Pro Val Pro Gly
```

```
    50                  55                  60
Glu Glu Gly Glu Leu Val Ser Thr Asp Pro Arg Pro Ala Ser Tyr Ser
65                  70                  75                  80
Phe Cys Ser Gly Lys Gly Val Gly Ile Lys Gly Glu Thr Ser Thr Ala
                85                  90                  95
Thr Pro Arg Arg Ser Asp Leu Asp Leu Gly Tyr Glu Pro Glu Gly Ser
            100                 105                 110
Ala Ser Pro Thr Pro Pro Tyr Leu Lys Trp Ala Glu Ser Leu His Ser
        115                 120                 125
Leu Leu Asp Asp Gln Asp Gly Ile Ser Leu Phe Arg Thr Phe Leu Lys
    130                 135                 140
Gln Glu Gly Cys Ala Asp Leu Leu Asp Phe Trp Phe Ala Cys Thr Gly
145                 150                 155                 160
Phe Arg Lys Leu Glu Pro Cys Asp Ser Asn Glu Glu Lys Arg Leu Lys
                165                 170                 175
Leu Ala Arg Ala Ile Tyr Arg Lys Tyr Ile Leu Asp Asn Asn Gly Ile
            180                 185                 190
Val Ser Arg Gln Thr Lys Pro Ala Thr Lys Ser Phe Ile Lys Gly Cys
        195                 200                 205
Ile Met Lys Gln Leu Ile Asp Pro Ala Met Phe Asp Gln Ala Gln Thr
    210                 215                 220
Glu Ile Gln Ala Thr Met Glu Glu Asn Thr Tyr Pro Ser Phe Leu Lys
225                 230                 235                 240
Ser Asp Ile Tyr Leu Glu Tyr Thr Arg Thr Gly Ser Glu Ser Pro Lys
                245                 250                 255
Val Cys Ser Asp Gln Ser Ser Gly Ser Gly Thr Gly Lys Gly Ile Ser
            260                 265                 270
Gly Tyr Leu Pro Thr Leu Asn Glu Asp Glu Glu Trp Lys Cys Asp Gln
        275                 280                 285
Asp Met Asp Glu Asp Asp Gly Arg Asp Ala Ala Pro Pro Gly Arg Leu
    290                 295                 300
Pro Gln Lys Leu Leu Leu Glu Thr Ala Ala Pro Arg Val Ser Ser Ser
305                 310                 315                 320
Arg Arg Tyr Ser Glu Gly Arg Glu Phe Arg Tyr Gly Ser Trp Arg Glu
                325                 330                 335
Pro Val Asn Pro Tyr Tyr Val Asn Ala Gly Tyr Ala Leu Ala Pro Ala
            340                 345                 350
Thr Ser Ala Asn Asp Ser Glu Gln Gln Ser Leu Ser Ser Asp Ala Asp
        355                 360                 365
Thr Leu Ser Leu Thr Asp Ser Ser Val Asp Gly Ile Pro Pro Tyr Arg
    370                 375                 380
Ile Arg Lys Gln His Arg Arg Glu Met Gln Glu Ser Ala Gln Val Asn
385                 390                 395                 400
Gly Arg Val Pro Leu Pro His Ile Pro Arg Thr Tyr Arg Val Pro Lys
                405                 410                 415
Glu Val Arg Val Glu Pro Gln Lys Phe Ala Glu Glu Leu Ile His Arg
            420                 425                 430
Leu Glu Ala Val Gln Arg Thr Arg Glu Ala Glu Glu Lys Leu Glu Glu
        435                 440                 445
Arg Leu Lys Arg Val Arg Met Glu Glu Glu Gly Glu Asp Gly Asp Pro
    450                 455                 460
Ser Ser Gly Pro Pro Gly Pro Cys His Lys Leu Pro Pro Ala Pro Ala
465                 470                 475                 480
Trp His His Phe Pro Pro Arg Leu Cys Trp Thr Trp Ala Cys Ala Gly
                485                 490                 495
Leu Arg Asp Ala His Glu Glu Asn Pro Glu Ser Ile Leu Asp Glu His
            500                 505                 510
```

Val Gln Arg Val Leu Arg Thr Thr Gly Arg Gln Ser Pro Gly Pro Gly
515 520 525
His Arg Ser Pro Asp Ser Gly His Val Ala Lys Met Pro Val Ala Leu
530 535 540
Gly Gly Ala Ala Ser Gly His Gly Lys His Val Pro Lys Ser Gly Ala
545 550 555 560
Lys Leu Asp Ala Ala Gly Leu His His His Arg His Val His His His
565 570 575
Val His His Ser Thr Ala Arg Pro Lys Glu Gln Val Glu Ala Glu Ala
580 585 590
Thr Arg Arg Ala Gln Ser Ser Phe Ala Trp Gly Leu Glu Pro His Ser
595 600 605
His Gly Ala Arg Ser Arg Gly Tyr Ser Glu Ser Val Gly Ala Ala Pro
610 615 620
Asn Ala Ser Asp Gly Leu Ala His Ser Gly Lys Val Gly Val Ala Cys
625 630 635 640
Lys Arg Asn Ala Lys Lys Ala Glu Ser Gly Lys Ser Ala Ser Thr Glu
645 650 655
Val Pro Gly Ala Ser Glu Asp Ala Glu Lys Asn Gln Lys Ile Met Gln
660 665 670
Trp Ile Ile Glu Gly Glu Lys Glu Ile Ser Arg His Arg Arg Thr Gly
675 680 685
His Gly Ser Ser Gly Thr Arg Lys Pro Gln Pro His Glu Asn Ser Arg
690 695 700
Pro Leu Ser Leu Glu His Pro Trp Ala Gly Pro Gln Leu Arg Thr Ser
705 710 715 720
Val Gln Pro Ser His Leu Phe Ile Gln Asp Pro Thr Met Pro Pro His
725 730 735
Pro Ala Pro Asn Pro Leu Thr Gln Leu Glu Glu Ala Arg Arg Arg Leu
740 745 750
Glu Glu Glu Glu Lys Arg Ala Ser Arg Ala Pro Ser Lys Gln Arg Tyr
755 760 765
Val Gln Glu Val Met Arg Arg Gly Arg Ala Cys Val Arg Pro Ala Cys
770 775 780
Ala Pro Val Leu His Val Val Pro Ala Val Ser Asp Met Glu Leu Ser
785 790 795 800
Glu Thr Glu Thr Arg Ser Gln Arg Lys Val Gly Gly Gly Ser Ala Gln
805 810 815
Pro Cys Asp Ser Ile Val Val Ala Tyr Tyr Phe Cys Gly Glu Pro Ile
820 825 830
Pro Tyr Arg Thr Leu Val Arg Gly Arg Ala Val Thr Leu Gly Gln Phe
835 840 845
Lys Glu Leu Leu Thr Lys Lys Gly Ser Tyr Arg Tyr Tyr Phe Lys Lys
850 855 860
Val Ser Asp Glu Phe Asp Cys Gly Val Val Phe Glu Glu Val Arg Glu
865 870 875 880
Asp Glu Ala Val Leu Pro Val Phe Glu Glu Lys Ile Ile Gly Lys Val
885 890 895
Glu Lys Val Asp
900

<210> 92
<211> 591
<212> PRT
<213> Homo sapiens

<400> 92

```
Met Val Pro Val Ala Val Thr Ala Ala Val Ala Pro Val Leu Ser Ile
1               5                   10                  15
Asn Ser Asp Phe Ser Asp Leu Arg Glu Ile Lys Lys Gln Leu Leu Leu
            20                  25                  30
Ile Ala Gly Leu Thr Arg Glu Arg Gly Leu Leu His Ser Ser Lys Trp
        35                  40                  45
Ser Ala Glu Leu Ala Phe Ser Leu Pro Ala Leu Pro Leu Ala Glu Leu
    50                  55                  60
Gln Pro Pro Pro Pro Ile Thr Glu Glu Asp Ala Gln Asp Met Asp Ala
65                  70                  75                  80
Tyr Thr Leu Ala Lys Ala Tyr Phe Asp Val Lys Glu Tyr Asp Arg Ala
                85                  90                  95
Ala His Phe Leu His Gly Cys Asn Ser Lys Lys Ala Tyr Phe Leu Tyr
            100                 105                 110
Met Tyr Ser Arg Tyr Leu Ser Gly Glu Lys Lys Lys Asp Asp Glu Thr
        115                 120                 125
Val Asp Ser Leu Gly Pro Leu Glu Lys Gly Gln Val Lys Asn Glu Ala
    130                 135                 140
Leu Arg Glu Leu Arg Val Glu Leu Ser Lys Lys His Gln Ala Arg Glu
145                 150                 155                 160
Leu Asp Gly Phe Gly Leu Tyr Leu Tyr Gly Val Val Leu Arg Lys Leu
                165                 170                 175
Asp Leu Val Lys Glu Ala Ile Asp Val Phe Val Glu Ala Thr His Val
            180                 185                 190
Leu Pro Leu His Trp Gly Ala Trp Leu Glu Leu Cys Asn Leu Ile Thr
        195                 200                 205
Asp Lys Glu Met Leu Lys Phe Leu Ser Leu Pro Asp Thr Trp Met Lys
    210                 215                 220
Glu Phe Phe Leu Ala His Ile Tyr Thr Glu Leu Gln Leu Ile Glu Glu
225                 230                 235                 240
Ala Leu Gln Lys Tyr Gln Asn Leu Ile Asp Val Gly Phe Ser Lys Ser
                245                 250                 255
Ser Tyr Ile Val Ser Gln Ile Ala Val Ala Tyr His Asn Ile Arg Asp
            260                 265                 270
Ile Asp Lys Ala Leu Ser Ile Phe Asn Glu Leu Arg Lys Gln Asp Pro
        275                 280                 285
Tyr Arg Ile Glu Asn Met Asp Thr Phe Ser Asn Leu Leu Tyr Val Arg
    290                 295                 300
Ser Met Lys Ser Glu Leu Ser Tyr Leu Ala His Asn Leu Cys Glu Ile
305                 310                 315                 320
Asp Lys Tyr Arg Val Glu Thr Cys Cys Val Ile Gly Asn Tyr Tyr Ser
                325                 330                 335
Leu Arg Ser Gln His Glu Lys Ala Ala Leu Tyr Phe Gln Arg Ala Leu
            340                 345                 350
Lys Leu Asn Pro Arg Tyr Leu Gly Ala Trp Thr Leu Met Gly His Glu
        355                 360                 365
Tyr Met Glu Met Lys Asn Thr Ser Ala Ala Ile Gln Ala Tyr Arg His
    370                 375                 380
Ala Ile Glu Val Asn Lys Arg Asp Tyr Arg Ala Trp Tyr Gly Leu Gly
385                 390                 395                 400
Gln Thr Tyr Glu Ile Leu Lys Met Pro Phe Tyr Cys Leu Tyr Tyr Tyr
                405                 410                 415
Arg Arg Ala His Gln Leu Arg Pro Asn Asp Ser Arg Met Leu Val Ala
            420                 425                 430
Leu Gly Glu Cys Tyr Glu Lys Leu Asn Gln Leu Val Glu Ala Lys Lys
        435                 440                 445
Cys Tyr Trp Arg Ala Tyr Ala Val Gly Asp Val Glu Lys Met Ala Leu
```

450 455 460
Val Lys Leu Ala Lys Leu His Glu Gln Leu Thr Glu Ser Glu Gln Ala
465 470 475 480
Ala Gln Cys Tyr Ile Lys Tyr Ile Gln Asp Ile Tyr Ser Cys Gly Glu
485 490 495
Ile Val Glu His Leu Glu Glu Ser Thr Ala Phe Arg Tyr Leu Ala Gln
500 505 510
Tyr Tyr Phe Lys Cys Lys Leu Trp Asp Glu Ala Ser Thr Cys Ala Gln
515 520 525
Lys Cys Cys Ala Phe Asn Asp Thr Arg Glu Glu Gly Lys Ala Leu Leu
530 535 540
Arg Gln Ile Leu Gln Leu Arg Asn Gln Gly Glu Thr Pro Thr Thr Glu
545 550 555 560
Val Pro Ala Pro Phe Phe Leu Pro Ala Ser Leu Ser Ala Asn Asn Thr
565 570 575
Pro Thr Arg Arg Val Ser Pro Leu Asn Leu Ser Ser Val Thr Pro
580 585 590

<210> 93
<211> 914
<212> PRT
<213> Homo sapiens

<400> 93
Val Tyr Gln Val Leu Leu Val Gly Ser Thr Leu Leu Lys Glu Val Pro
1 5 10 15
Ser Gly Leu Gln Leu Glu Gln Leu Pro Ser Gln Ser Leu Leu Thr His
20 25 30
Ile Pro Thr Ala Gly Leu Pro Thr Ser Leu Gly Gly Gly Leu Pro Tyr
35 40 45
Cys His Gln Ala Trp Leu Asp Phe Arg Arg Arg Leu Glu Ala Leu Leu
50 55 60
Gln Asn Cys Gln Ala Ala Cys Ala Leu Leu Gln Gly Ala Ile Glu Ser
65 70 75 80
Val Lys Ala Val Pro Gln Pro Met Glu Pro Gly Glu Val Gly Gln Leu
85 90 95
Leu Gln Gln Thr Glu Val Leu Met Gln Gln Val Leu Asp Ser Pro Trp
100 105 110
Leu Ala Trp Leu Gln Cys Gln Gly Gly Arg Glu Leu Thr Trp Leu Lys
115 120 125
Gln Glu Val Pro Glu Val Thr Leu Ser Pro Asp Tyr Arg Thr Ala Met
130 135 140
Asp Lys Ala Asp Glu Leu Tyr Asp Arg Val Asp Gly Leu Leu His Gln
145 150 155 160
Leu Thr Leu Gln Ser Asn Gln Arg Ile Gln Ala Leu Glu Leu Val Gln
165 170 175
Thr Leu Glu Ala Arg Glu Ser Gly Leu His Gln Ile Glu Val Trp Leu
180 185 190
Gln Gln Val Gly Trp Pro Ala Leu Glu Glu Ala Gly Glu Pro Ser Leu
195 200 205
Asp Met Leu Leu Gln Ala Gln Gly Ser Phe Gln Glu Leu Tyr Gln Val
210 215 220
Ala Gln Glu Gln Val Arg Gln Gly Glu Lys Phe Leu Gln Pro Leu Thr
225 230 235 240
Gly Trp Glu Ala Ala Glu Leu Asp Pro Pro Gly Ala Arg Phe Leu Ala
245 250 255
Leu Arg Ala Gln Leu Thr Glu Phe Ser Arg Ala Leu Ala Gln Arg Cys

```
                260                 265                 270
Gln Arg Leu Ala Asp Ala Glu Arg Leu Phe Gln Leu Phe Arg Glu Ala
        275                 280                 285
Leu Thr Trp Ala Glu Glu Gly Gln Arg Val Leu Ala Glu Leu Glu Gln
    290                 295                 300
Glu Arg Pro Gly Val Val Leu Gln Gln Leu Gln Leu His Trp Thr Arg
305                 310                 315                 320
His Pro Asp Leu Pro Pro Ala His Phe Arg Lys Met Trp Ala Leu Ala
                325                 330                 335
Thr Gly Leu Gly Ser Glu Ala Ile Arg Gln Glu Cys Arg Trp Ala Trp
            340                 345                 350
Ala Arg Cys Gln Asp Thr Trp Leu Ala Leu Asp Gln Lys Leu Glu Ala
        355                 360                 365
Ser Leu Lys Leu Pro Pro Val Gly Ser Thr Ala Ser Leu Cys Val Ser
    370                 375                 380
Gln Val Pro Ala Ala Pro Ala His Pro Pro Leu Arg Lys Ala Tyr Ser
385                 390                 395                 400
Phe Asp Arg Asn Leu Gly Gln Ser Leu Ser Glu Pro Ala Cys His Cys
                405                 410                 415
His His Ala Ala Thr Ile Ala Ala Cys Arg Arg Pro Glu Ala Gly Gly
            420                 425                 430
Gly Ala Leu Pro Gln Ala Ser Pro Thr Val Pro Pro Pro Gly Ser Ser
        435                 440                 445
Asp Pro Arg Ser Leu Asn Arg Leu Gln Leu Val Leu Ala Glu Met Val
    450                 455                 460
Ala Thr Glu Arg Glu Tyr Val Arg Ala Leu Glu Tyr Thr Met Glu Asn
465                 470                 475                 480
Tyr Phe Pro Glu Leu Asp Arg Pro Asp Val Pro Gln Gly Leu Arg Gly
                485                 490                 495
Gln Arg Ala His Leu Phe Gly Asn Leu Glu Lys Leu Arg Asp Phe His
            500                 505                 510
Cys His Phe Phe Leu Arg Glu Leu Glu Ala Cys Thr Arg His Pro Pro
        515                 520                 525
Arg Val Ala Tyr Ala Phe Leu Arg His Arg Val Gln Phe Gly Met Tyr
    530                 535                 540
Ala Leu Tyr Ser Lys Asn Lys Pro Arg Ser Asp Ala Leu Met Ser Ser
545                 550                 555                 560
Tyr Gly His Thr Phe Phe Lys Asp Lys Gln Gln Ala Leu Gly Asp His
                565                 570                 575
Leu Asp Leu Ala Ser Tyr Leu Leu Lys Pro Ile Gln Arg Met Gly Lys
            580                 585                 590
Tyr Ala Leu Leu Leu Gln Glu Leu Ala Arg Ala Cys Gly Gly Pro Thr
        595                 600                 605
Gln Glu Leu Ser Ala Leu Arg Glu Ala Gln Ser Leu Val His Phe Gln
    610                 615                 620
Leu Arg His Gly Asn Asp Leu Leu Ala Met Asp Ala Ile Gln Gly Cys
625                 630                 635                 640
Asp Val Asn Leu Lys Glu Gln Gly Gln Leu Val Arg Gln Asp Glu Phe
                645                 650                 655
Val Val Arg Thr Gly Arg His Lys Ser Val Arg Arg Ile Phe Leu Phe
            660                 665                 670
Glu Glu Leu Leu Leu Phe Ser Lys Pro Arg His Gly Pro Thr Gly Val
        675                 680                 685
Asp Thr Phe Ala Tyr Lys Arg Ser Phe Lys Met Ala Asp Leu Gly Leu
    690                 695                 700
Thr Glu Cys Cys Gly Asn Ser Asn Leu Arg Phe Glu Ile Trp Phe Arg
705                 710                 715                 720
```

```
Arg Arg Lys Ala Arg Asp Thr Phe Val Leu Gln Ala Ser Ser Leu Ala
                725                 730                 735
Ile Lys Gln Ala Trp Thr Ala Asp Ile Ser His Leu Leu Trp Arg Gln
            740                 745                 750
Ala Val His Asn Lys Glu Val Arg Met Ala Glu Met Val Ser Met Gly
        755                 760                 765
Val Gly Asn Lys Ala Phe Arg Asp Ile Ala Pro Ser Glu Glu Ala Ile
    770                 775                 780
Asn Asp Arg Thr Val Asn Tyr Val Leu Lys Cys Arg Glu Val Arg Ser
785                 790                 795                 800
Arg Ala Ser Ile Ala Val Ala Pro Phe Asp His Asp Ser Leu Tyr Leu
                805                 810                 815
Gly Ala Ser Asn Ser Leu Pro Gly Asp Pro Ala Ser Cys Ser Val Leu
            820                 825                 830
Gly Ser Leu Asn Leu His Leu Tyr Arg Asp Pro Ala Leu Leu Gly Leu
        835                 840                 845
Arg Cys Pro Leu Tyr Pro Ser Phe Leu Glu Glu Ala Ala Leu Glu Ala
    850                 855                 860
Glu Ala Glu Leu Gly Gly Gln Pro Ser Leu Thr Ala Glu Asp Ser Glu
865                 870                 875                 880
Ile Ser Ser Gln Cys Pro Ser Ala Ser Gly Ser Ser Gly Ser Asp Ser
                885                 890                 895
Ser Cys Val Ser Gly Gln Ala Leu Gly Arg Gly Leu Glu Asp Leu Pro
            900                 905                 910
Cys Val


&lt;210&gt; 94
&lt;211&gt; 277
&lt;212&gt; PRT
&lt;213&gt; Homo sapiens

&lt;400&gt; 94
Leu Asn Tyr Leu Leu Glu Ser Arg Leu Glu Ala Ala Ala His Cys Ala
1               5                   10                  15
Leu Lys Gln Gly Ile Ala Thr Ala Ser Leu Leu Pro Ala Gln Leu Gln
            20                  25                  30
Pro Ala Val Leu Thr Val Val Thr Cys His Val Val Val Ser Val His
        35                  40                  45
Gly His His Thr Asp Gly Cys Leu Ala Ala Leu Cys Arg Glu Asp Arg
    50                  55                  60
Thr Gly Thr Gly Gly Ala Phe Trp Cys Lys Asn Arg Val Ile Val Ser
65                  70                  75                  80
His Ala Val Asp Val Val Leu His Val His Gly Glu Gly Asn Pro Val
                85                  90                  95
Gln Ala Leu Ile Ala His Gly Ala Pro Glu Ala Ala Trp Val Val Gly
            100                 105                 110
Leu Ala Gln Gly Leu Gln Asp His Phe His Asp Glu Met Ser Thr His
        115                 120                 125
Ala Ala Phe Val Gly Arg Leu Leu Glu Pro Gly Val Gln Glu Val Leu
    130                 135                 140
Leu Ala Val His Phe Leu Thr His Val Val Glu Arg Leu Pro Thr Glu
145                 150                 155                 160
Ser Ser Pro Thr Arg Val Ala Gly Glu Ala Val Ser Val Ile Lys Thr
                165                 170                 175
Pro His Cys Leu Ala Arg Leu Leu Gly Ser Val Asp Ala Lys Pro Thr
            180                 185                 190
```

Leu Asp Ala Asn Ala Glu Val Val Pro Arg Arg Ala Arg Leu Glu Arg
195 200 205
Pro Leu Gln Leu Pro Gly Glu Arg Leu Gln Pro Pro Leu Gly Arg Ala
210 215 220
Trp Ala Ala Leu Pro Ala Arg Gly Gln Arg Glu Cys Arg Gln Arg Glu
225 230 235 240
Gly Gly Arg Pro Arg Arg Leu Arg Gly Ala Ser Gly Arg Gly Ala Gly
245 250 255
Ala Gly Arg Glu Glu Val Ser Val Gly Phe Ser Ala Gln Trp Glu Phe
260 265 270
Gly Ser Gly Arg His
275

<210> 95
<211> 1120
<212> PRT
<213> Homo sapiens

<400> 95
Met Trp Arg Val Lys Lys Leu Ser Leu Ser Leu Ser Pro Ser Pro Gln
1 5 10 15
Thr Gly Lys Pro Ser Met Arg Thr Pro Leu Arg Glu Leu Thr Leu Gln
20 25 30
Pro Gly Ala Leu Thr Thr Ser Gly Lys Arg Ser Pro Ala Cys Ser Ser
35 40 45
Leu Thr Pro Ser Leu Cys Lys Leu Gly Leu Gln Glu Gly Ser Asn Asn
50 55 60
Ser Ser Pro Val Asp Phe Val Asn Asn Lys Arg Thr Asp Leu Ser Ser
65 70 75 80
Glu His Phe Ser His Ser Ser Lys Trp Leu Glu Thr Cys Gln His Glu
85 90 95
Ser Asp Glu Gln Pro Leu Asp Pro Ile Pro Gln Ile Ser Ser Thr Pro
100 105 110
Lys Thr Ser Glu Glu Ala Val Asp Pro Leu Gly Asn Tyr Met Val Lys
115 120 125
Thr Ile Val Leu Val Pro Ser Pro Leu Gly Gln Gln Asp Met Ile
130 135 140
Phe Glu Ala Arg Leu Asp Thr Met Ala Glu Thr Asn Ser Ile Ser Leu
145 150 155 160
Asn Gly Pro Leu Arg Thr Asp Asp Leu Val Arg Glu Glu Val Ala Pro
165 170 175
Cys Met Gly Asp Arg Phe Ser Glu Val Ala Ala Val Ser Glu Lys Pro
180 185 190
Ile Phe Gln Glu Ser Pro Ser His Leu Leu Glu Glu Ser Pro Pro Asn
195 200 205
Pro Cys Ser Glu Gln Leu His Cys Ser Lys Glu Ser Leu Ser Ser Arg
210 215 220
Thr Glu Ala Val Arg Glu Asp Leu Val Pro Ser Glu Ser Asn Ala Phe
225 230 235 240
Leu Pro Ser Ser Val Leu Trp Leu Ser Pro Ser Thr Ala Leu Ala Ala
245 250 255
Asp Phe Arg Val Asn His Val Asp Pro Glu Glu Glu Ile Val Glu His
260 265 270
Gly Ala Met Glu Glu Arg Glu Met Arg Phe Pro Thr His Pro Lys Glu
275 280 285
Ser Glu Thr Glu Asp Gln Ala Leu Val Ser Ser Val Glu Asp Ile Leu
290 295 300

Ser Thr Cys Leu Thr Pro Asn Leu Val Glu Met Glu Ser Gln Glu Ala
305 310 315 320
Pro Gly Pro Ala Val Glu Asp Val Gly Arg Ile Leu Gly Ser Asp Thr
325 330 335
Glu Ser Trp Met Ser Pro Leu Ala Trp Leu Glu Lys Gly Val Asn Thr
340 345 350
Ser Val Met Leu Glu Asn Leu Arg Gln Ser Leu Ser Leu Pro Ser Met
355 360 365
Leu Arg Asp Ala Ala Ile Gly Thr Thr Pro Phe Ser Thr Cys Ser Val
370 375 380
Gly Thr Trp Phe Thr Pro Ser Ala Pro Gln Glu Lys Ser Thr Asn Thr
385 390 395 400
Ser Gln Thr Gly Leu Val Gly Thr Lys His Ser Thr Ser Glu Thr Glu
405 410 415
Gln Leu Leu Cys Gly Arg Pro Pro Asp Leu Thr Ala Leu Ser Arg His
420 425 430
Asp Leu Glu Asp Asn Leu Leu Ser Ser Leu Val Ile Val Glu Phe Leu
435 440 445
Ser Arg Gln Leu Arg Asp Trp Lys Ser Gln Leu Ala Val Pro His Pro
450 455 460
Glu Thr Gln Asp Ser Ser Thr Gln Thr Asp Thr Ser His Ser Gly Ile
465 470 475 480
Thr Asn Lys Leu Gln His Leu Lys Glu Ser His Glu Met Gly Gln Ala
485 490 495
Leu Gln Gln Ala Arg Asn Val Met Gln Ser Trp Val Leu Ile Ser Lys
500 505 510
Glu Leu Ile Ser Leu Leu His Leu Ser Leu Leu His Leu Glu Glu Asp
515 520 525
Lys Thr Thr Val Asn Gln Glu Ser Arg Arg Ala Glu Thr Leu Val Cys
530 535 540
Cys Cys Phe Asp Leu Leu Lys Lys Leu Arg Ala Lys Leu Gln Ser Leu
545 550 555 560
Lys Ala Glu Arg Glu Glu Ala Arg His Arg Glu Glu Met Ala Leu Arg
565 570 575
Gly Lys Asp Ala Ala Glu Ile Val Leu Glu Ala Phe Cys Ala His Ala
580 585 590
Ser Gln Arg Ile Ser Gln Leu Glu Gln Asp Leu Ala Ser Met Arg Glu
595 600 605
Phe Arg Gly Leu Leu Lys Asp Ala Gln Thr Gln Leu Val Gly Leu His
610 615 620
Ala Lys Gln Glu Glu Leu Val Gln Gln Thr Val Ser Leu Thr Ser Thr
625 630 635 640
Leu Gln Gln Asp Trp Arg Ser Met Gln Leu Asp Tyr Thr Thr Trp Thr
645 650 655
Ala Leu Leu Ser Arg Ser Arg Gln Leu Thr Glu Lys Leu Thr Val Lys
660 665 670
Ser Gln Gln Ala Leu Gln Glu Arg Asp Val Ala Ile Glu Glu Lys Gln
675 680 685
Glu Val Ser Arg Val Leu Glu Gln Val Ser Ala Gln Leu Glu Glu Cys
690 695 700
Lys Gly Gln Thr Glu Gln Leu Glu Leu Glu Asn Ile Arg Leu Ala Thr
705 710 715 720
Asp Leu Arg Ala Gln Leu Gln Ile Leu Ala Asn Met Asp Ser Gln Leu
725 730 735
Lys Glu Leu Gln Ser Gln His Thr His Cys Ala Gln Asp Leu Ala Met
740 745 750
Lys Asp Glu Leu Leu Cys Gln Leu Thr Gln Ser Asn Glu Glu Gln Ala

755 760 765
Ala Gln Cys Val Lys Glu Glu Met Ala Leu Lys His Met Gln Ala Glu
770 775 780
Leu Gln Gln Gln Gln Ala Val Leu Ala Lys Glu Val Arg Asp Leu Lys
785 790 795 800
Glu Thr Leu Glu Phe Ala Asp Gln Glu Asn Gln Val Ala His Leu Glu
805 810 815
Leu Gly Gln Val Glu Cys Gln Leu Lys Thr Thr Leu Glu Val Leu Arg
820 825 830
Glu Arg Ser Leu Gln Cys Glu Asn Leu Lys Asp Thr Val Glu Asn Leu
835 840 845
Thr Ala Lys Leu Ala Ser Thr Ile Ala Asp Asn Gln Glu Gln Asp Leu
850 855 860
Glu Lys Thr Arg Gln Tyr Ser Gln Lys Leu Gly Leu Leu Thr Glu Gln
865 870 875 880
Leu Gln Ser Leu Thr Leu Phe Leu Gln Thr Lys Leu Lys Glu Lys Thr
885 890 895
Glu Gln Glu Thr Leu Leu Leu Ser Thr Ala Cys Pro Pro Thr Gln Glu
900 905 910
His Pro Leu Pro Asn Asp Arg Thr Phe Leu Gly Ser Ile Leu Thr Ala
915 920 925
Val Ala Asp Glu Glu Pro Glu Ser Thr Pro Val Pro Leu Leu Gly Ser
930 935 940
Asp Lys Ser Ala Phe Thr Arg Val Ala Ser Met Val Ser Leu Gln Pro
945 950 955 960
Ala Glu Thr Pro Gly Met Glu Glu Ser Leu Ala Glu Met Ser Ile Met
965 970 975
Thr Thr Glu Leu Gln Ser Leu Cys Ser Leu Leu Gln Glu Ser Lys Glu
980 985 990
Glu Ala Ile Arg Thr Leu Gln Arg Lys Ile Cys Glu Leu Gln Ala Arg
995 1000 1005
Leu Gln Ala Gln Glu Glu Gln His Gln Glu Val Gln Lys Ala Lys Glu
1010 1015 1020
Ala Asp Ile Glu Lys Leu Asn Gln Ala Leu Cys Leu Arg Tyr Lys Asn
1025 1030 1035 1040
Glu Lys Glu Leu Gln Glu Val Ile Gln Gln Asn Glu Lys Ile Leu Glu
1045 1050 1055
Gln Ile Asp Lys Ser Gly Glu Leu Ile Ser Leu Arg Glu Glu Val Thr
1060 1065 1070
His Leu Thr Arg Ser Leu Arg Arg Ala Glu Thr Glu Thr Lys Val Leu
1075 1080 1085
Gln Glu Ala Trp Gln Ala Ser Trp Thr Pro Thr Ala Ser Leu Trp Pro
1090 1095 1100
Pro Ile Gly Ser Arg Arg Lys Cys Gly Ser Leu Arg Arg Trp Thr Asn
1105 1110 1115 1120

<210> 96
<211> 540
<212> PRT
<213> Homo sapiens

<400> 96
Met Gly Thr Thr Ala Arg Ala Ala Leu Val Leu Thr Tyr Leu Ala Val
1 5 10 15
Ala Ser Ala Ala Ser Glu Gly Gly Phe Thr Ala Thr Gly Gln Arg Gln
20 25 30
Leu Arg Pro Glu His Phe Gln Glu Val Gly Tyr Ala Ala Pro Pro Ser

```
        35                      40                      45
Pro Pro Leu Ser Arg Ser Leu Pro Met Asp His Pro Asp Ser Ser Gln
    50                      55                      60
His Gly Pro Pro Phe Glu Gly Gln Ser Gln Val Gln Pro Pro Pro Ser
65                      70                      75                  80
Gln Glu Ala Thr Pro Leu Gln Gln Glu Lys Leu Leu Pro Ala Gln Leu
                85                      90                      95
Pro Ala Glu Lys Glu Val Gly Pro Pro Leu Pro Gln Glu Ala Val Pro
            100                     105                     110
Leu Gln Lys Glu Leu Pro Ser Leu Gln His Pro Asn Glu Gln Lys Glu
        115                     120                     125
Gly Thr Pro Ala Pro Phe Gly Asp Gln Ser His Pro Glu Pro Glu Ser
    130                     135                     140
Trp Asn Ala Ala Gln His Cys Gln Gln Asp Arg Ser Gln Gly Gly Trp
145                     150                     155                 160
Gly His Arg Leu Asp Gly Phe Pro Pro Gly Arg Pro Ser Pro Asp Asn
                165                     170                     175
Leu Asn Gln Ile Cys Leu Pro Asn Arg Gln His Val Val Tyr Gly Pro
            180                     185                     190
Trp Asn Leu Pro Gln Ser Ser Tyr Ser His Leu Thr Arg Gln Gly Glu
        195                     200                     205
Thr Leu Asn Phe Leu Glu Ile Gly Tyr Ser Arg Cys Cys His Cys Arg
    210                     215                     220
Ser His Thr Asn Arg Leu Glu Cys Ala Lys Leu Val Trp Glu Glu Ala
225                     230                     235                 240
Met Ser Arg Phe Cys Glu Ala Glu Phe Ser Val Lys Thr Arg Pro His
                245                     250                     255
Trp Cys Cys Thr Arg Gln Gly Glu Ala Arg Phe Ser Cys Phe Gln Glu
            260                     265                     270
Glu Ala Pro Gln Pro His Tyr Gln Leu Arg Ala Cys Pro Ser His Gln
        275                     280                     285
Pro Asp Ile Ser Ser Gly Leu Glu Leu Pro Phe Pro Pro Gly Val Pro
    290                     295                     300
Thr Leu Asp Asn Ile Lys Asn Ile Cys His Leu Arg Arg Phe Arg Ser
305                     310                     315                 320
Val Pro Arg Asn Leu Pro Ala Thr Asp Pro Leu Gln Arg Glu Leu Leu
                325                     330                     335
Ala Leu Ile Gln Leu Glu Arg Glu Phe Gln Arg Cys Cys Arg Gln Gly
            340                     345                     350
Asn Asn His Thr Cys Thr Trp Lys Ala Trp Glu Asp Thr Leu Asp Lys
        355                     360                     365
Tyr Cys Asp Arg Glu Tyr Ala Val Lys Thr His His His Leu Cys Cys
    370                     375                     380
Arg His Pro Pro Ser Pro Thr Arg Asp Glu Cys Phe Ala Arg Arg Ala
385                     390                     395                 400
Pro Tyr Pro Asn Tyr Asp Arg Asp Ile Leu Thr Ile Asp Ile Ser Arg
                405                     410                     415
Val Thr Pro Asn Leu Met Gly His Leu Cys Gly Asn Gln Arg Val Leu
            420                     425                     430
Thr Lys His Lys His Ile Pro Gly Leu Ile His Asn Met Thr Ala Arg
        435                     440                     445
Cys Cys Asp Leu Pro Phe Pro Glu Gln Ala Cys Cys Ala Glu Glu Glu
    450                     455                     460
Lys Leu Thr Phe Ile Asn Asp Leu Cys Gly Pro Arg Arg Asn Ile Trp
465                     470                     475                 480
Arg Asp Pro Ala Leu Cys Cys Tyr Leu Ser Pro Gly Asp Glu Gln Val
                485                     490                     495
```

Asn Cys Phe Asn Ile Asn Tyr Leu Arg Asn Val Ala Leu Val Ser Gly
500 505 510
Asp Thr Glu Asn Ala Lys Gly Gln Gly Glu Gln Gly Ser Thr Gly Gly
515 520 525
Thr Asn Ile Ser Ser Thr Ser Glu Pro Lys Glu Glu
530 535 540

<210> 97
<211> 462
<212> PRT
<213> Homo sapiens

<400> 97
Met Gly Lys Glu Lys Thr His Ile Asn Ile Val Val Ile Gly His Val
1 5 10 15
Asp Ser Gly Lys Ser Thr Thr Thr Gly His Leu Ile Tyr Lys Cys Gly
20 25 30
Gly Ile Asp Lys Arg Thr Ile Glu Lys Phe Glu Lys Glu Ala Ala Glu
35 40 45
Met Gly Lys Gly Ser Phe Lys Tyr Ala Trp Val Leu Asp Lys Leu Lys
50 55 60
Ala Glu Arg Glu Arg Gly Ile Thr Ile Asp Ile Ser Leu Trp Lys Phe
65 70 75 80
Glu Thr Ser Lys Tyr Tyr Val Thr Ile Ile Asp Ala Pro Gly His Arg
85 90 95
Asp Phe Ile Lys Asn Met Ile Thr Gly Thr Ser Gln Ala Asp Cys Ala
100 105 110
Val Leu Ile Val Ala Ala Gly Val Gly Glu Phe Glu Ala Gly Ile Ser
115 120 125
Lys Asn Gly Gln Thr Arg Glu His Ala Leu Leu Ala Tyr Thr Leu Gly
130 135 140
Val Lys Gln Leu Ile Val Gly Val Asn Lys Met Asp Ser Thr Glu Pro
145 150 155 160
Pro Tyr Ser Gln Lys Arg Tyr Glu Glu Ile Val Lys Glu Val Ser Thr
165 170 175
Tyr Ile Lys Lys Ile Gly Tyr Asn Pro Asp Thr Val Ala Phe Val Pro
180 185 190
Ile Ser Gly Trp Asn Gly Asp Asn Met Leu Glu Pro Ser Ala Asn Met
195 200 205
Pro Trp Phe Lys Gly Trp Lys Val Thr Arg Lys Asp Gly Asn Ala Ser
210 215 220
Gly Thr Thr Leu Leu Glu Ala Val Asp Cys Ile Leu Pro Pro Thr Arg
225 230 235 240
Pro Thr Asp Lys Pro Leu Arg Leu Pro Leu Gln Asp Val Tyr Lys Ile
245 250 255
Gly Gly Ile Gly Thr Val Pro Val Gly Arg Val Glu Thr Gly Val Leu
260 265 270
Lys Pro Gly Met Val Val Thr Phe Ala Pro Val Asn Val Thr Thr Glu
275 280 285
Val Lys Ser Val Glu Met His His Glu Ala Leu Ser Glu Ala Leu Pro
290 295 300
Gly Asp Asn Val Gly Phe Asn Val Lys Asn Val Ser Val Lys Asp Val
305 310 315 320
Arg Arg Gly Asn Val Ala Gly Asp Ser Lys Asn Asp Pro Pro Met Glu
325 330 335
Ala Ala Gly Phe Thr Ala Gln Val Ile Ile Leu Asn His Pro Gly Gln
340 345 350

Ile Ser Ala Gly Tyr Ala Pro Val Leu Asp Cys His Thr Ala His Ile
355 360 365
Ala Cys Lys Phe Ala Glu Leu Lys Glu Lys Ile Asp Arg Arg Ser Gly
370 375 380
Lys Lys Leu Glu Asp Gly Pro Lys Phe Leu Lys Ser Gly Asp Ala Ala
385 390 395 400
Ile Val Asp Met Val Pro Gly Lys Pro Met Cys Val Glu Ser Phe Ser
405 410 415
Asp Tyr Pro Pro Leu Gly Arg Phe Ala Val Arg Asp Met Arg Gln Thr
420 425 430
Val Ala Val Gly Val Ile Lys Ala Val Asp Lys Lys Ala Ala Gly Ala
435 440 445
Gly Lys Val Thr Lys Ser Ala Gln Lys Ala Gln Lys Ala Lys
450 455 460

<210> 98
<211> 2328
<212> PRT
<213> Homo sapiens

<400> 98
Lys Ser Lys Arg Gln Ala Gln Gln Met Val Gln Pro Gln Ser Pro Val
1 5 10 15
Ala Val Ser Gln Ser Lys Pro Gly Cys Tyr Asp Asn Gly Lys His Tyr
20 25 30
Gln Ile Asn Gln Gln Trp Glu Arg Thr Tyr Leu Gly Asn Val Leu Val
35 40 45
Cys Thr Cys Tyr Gly Gly Ser Arg Gly Phe Asn Cys Glu Ser Lys Pro
50 55 60
Glu Ala Glu Glu Thr Cys Phe Asp Lys Tyr Thr Gly Asn Thr Tyr Arg
65 70 75 80
Val Gly Asp Thr Tyr Glu Arg Pro Lys Asp Ser Met Ile Trp Asp Cys
85 90 95
Thr Cys Ile Gly Ala Gly Arg Gly Arg Ile Ser Cys Thr Ile Ala Asn
100 105 110
Arg Cys His Glu Gly Gly Gln Ser Tyr Lys Ile Gly Asp Thr Trp Arg
115 120 125
Arg Pro His Glu Thr Gly Gly Tyr Met Leu Glu Cys Val Cys Leu Gly
130 135 140
Asn Gly Lys Gly Glu Trp Thr Cys Lys Pro Ile Ala Glu Lys Cys Phe
145 150 155 160
Asp His Ala Ala Gly Thr Ser Tyr Val Val Gly Glu Thr Trp Glu Lys
165 170 175
Pro Tyr Gln Gly Trp Met Met Val Asp Cys Thr Cys Leu Gly Glu Gly
180 185 190
Ser Gly Arg Ile Thr Cys Thr Ser Arg Asn Arg Cys Asn Asp Gln Asp
195 200 205
Thr Arg Thr Ser Tyr Arg Ile Gly Asp Thr Trp Ser Lys Lys Asp Asn
210 215 220
Arg Gly Asn Leu Leu Gln Cys Ile Cys Thr Gly Asn Gly Arg Gly Glu
225 230 235 240
Trp Lys Cys Glu Arg His Thr Ser Val Gln Thr Thr Ser Ser Gly Ser
245 250 255
Gly Pro Phe Thr Asp Val Arg Ala Ala Val Tyr Gln Pro Gln Pro His
260 265 270
Pro Gln Pro Pro Pro Tyr Gly His Cys Val Thr Asp Ser Gly Val Val
275 280 285

Tyr Ser Val Gly Met Gln Trp Leu Lys Thr Gln Gly Asn Lys Gln Met
290 295 300
Leu Cys Thr Cys Leu Gly Asn Gly Val Ser Cys Gln Glu Thr Ala Val
305 310 315 320
Thr Gln Thr Tyr Gly Gly Asn Leu Asn Gly Glu Pro Cys Val Leu Pro
325 330 335
Phe Thr Tyr Asn Gly Arg Thr Phe Tyr Ser Cys Thr Thr Glu Gly Arg
340 345 350
Gln Asp Gly His Leu Trp Cys Ser Thr Thr Ser Asn Tyr Glu Gln Asp
355 360 365
Gln Lys Tyr Ser Phe Cys Thr Asp His Thr Val Leu Val Gln Thr Gln
370 375 380
Gly Gly Asn Ser Asn Gly Ala Leu Cys His Phe Pro Phe Leu Tyr Asn
385 390 395 400
Asn His Asn Tyr Thr Asp Cys Thr Ser Glu Gly Arg Arg Asp Asn Met
405 410 415
Lys Trp Cys Gly Thr Thr Gln Asn Tyr Asp Ala Asp Gln Lys Phe Gly
420 425 430
Phe Cys Pro Met Ala Ala His Glu Glu Ile Cys Thr Thr Asn Glu Gly
435 440 445
Val Met Tyr Arg Ile Gly Asp Gln Trp Asp Lys Gln His Asp Met Gly
450 455 460
His Met Met Arg Cys Thr Cys Val Gly Asn Gly Arg Gly Glu Trp Thr
465 470 475 480
Cys Ile Ala Tyr Ser Gln Leu Arg Asp Gln Cys Ile Val Asp Asp Ile
485 490 495
Thr Tyr Asn Val Asn Asp Thr Phe His Lys Arg His Glu Glu Gly His
500 505 510
Met Leu Asn Cys Thr Cys Phe Gly Gln Gly Arg Gly Arg Trp Lys Cys
515 520 525
Asp Pro Val Asp Gln Cys Gln Asp Ser Glu Thr Gly Thr Phe Tyr Gln
530 535 540
Ile Gly Asp Ser Trp Glu Lys Tyr Val His Gly Val Arg Tyr Gln Cys
545 550 555 560
Tyr Cys Tyr Gly Arg Gly Ile Gly Glu Trp His Cys Gln Pro Leu Gln
565 570 575
Thr Tyr Pro Ser Ser Ser Gly Pro Val Glu Val Phe Ile Thr Glu Thr
580 585 590
Pro Ser Gln Pro Asn Ser His Pro Ile Gln Trp Asn Ala Pro Gln Pro
595 600 605
Ser His Ile Ser Lys Tyr Ile Leu Arg Trp Arg Pro Lys Asn Ser Val
610 615 620
Gly Arg Trp Lys Glu Ala Thr Ile Pro Gly His Leu Asn Ser Tyr Thr
625 630 635 640
Ile Lys Gly Leu Lys Pro Gly Val Val Tyr Glu Gly Gln Leu Ile Ser
645 650 655
Ile Gln Gln Tyr Gly His Gln Glu Val Thr Arg Phe Asp Phe Thr Thr
660 665 670
Thr Ser Thr Ser Thr Pro Val Thr Ser Asn Thr Val Thr Gly Glu Thr
675 680 685
Thr Pro Phe Ser Pro Leu Val Ala Thr Ser Glu Ser Val Thr Glu Ile
690 695 700
Thr Ala Ser Ser Phe Val Val Ser Trp Val Ser Ala Ser Asp Thr Val
705 710 715 720
Ser Gly Phe Arg Val Glu Tyr Glu Leu Ser Glu Glu Gly Asp Glu Pro
725 730 735
Gln Tyr Leu Asp Leu Pro Ser Thr Ala Thr Ser Val Asn Ile Pro Asp

```
            740                 745                 750
Leu Leu Pro Gly Arg Lys Tyr Ile Val Asn Val Tyr Gln Ile Ser Glu
        755                 760                 765
Asp Gly Glu Gln Ser Leu Ile Leu Ser Thr Ser Gln Thr Thr Ala Pro
    770                 775                 780
Asp Ala Pro Pro Asp Pro Thr Val Asp Gln Val Asp Asp Thr Ser Ile
785                 790                 795                 800
Val Val Arg Trp Ser Arg Pro Gln Ala Pro Ile Thr Gly Tyr Arg Ile
                805                 810                 815
Val Tyr Ser Pro Ser Val Glu Gly Ser Ser Thr Glu Leu Asn Leu Pro
            820                 825                 830
Glu Thr Ala Asn Ser Val Thr Leu Ser Asp Leu Gln Pro Gly Val Gln
        835                 840                 845
Tyr Asn Ile Thr Ile Tyr Ala Val Glu Glu Asn Gln Glu Ser Thr Pro
    850                 855                 860
Val Val Ile Gln Gln Glu Thr Thr Gly Thr Pro Arg Ser Asp Thr Val
865                 870                 875                 880
Pro Ser Pro Arg Asp Leu Gln Phe Val Glu Val Thr Asp Val Lys Val
                885                 890                 895
Thr Ile Met Trp Thr Pro Pro Glu Ser Ala Val Thr Gly Tyr Arg Val
            900                 905                 910
Asp Val Ile Pro Val Asn Leu Pro Gly Glu His Gly Gln Arg Leu Pro
        915                 920                 925
Ile Ser Arg Asn Thr Phe Ala Glu Val Thr Gly Leu Ser Pro Gly Val
    930                 935                 940
Thr Tyr Tyr Phe Lys Val Phe Ala Val Ser His Gly Arg Glu Ser Lys
945                 950                 955                 960
Pro Leu Thr Ala Gln Gln Thr Thr Lys Leu Asp Ala Pro Thr Asn Leu
                965                 970                 975
Gln Phe Val Asn Glu Thr Asp Ser Thr Val Leu Val Arg Trp Thr Pro
            980                 985                 990
Pro Arg Ala Gln Ile Thr Gly Tyr Arg Leu Thr Val Gly Leu Thr Arg
        995                 1000                1005
Arg Gly Gln Pro Arg Gln Tyr Asn Val Gly Pro Ser Val Ser Lys Tyr
    1010                1015                1020
Pro Leu Arg Asn Leu Gln Pro Ala Ser Glu Tyr Thr Val Ser Leu Val
1025                1030                1035                1040
Ala Ile Lys Gly Asn Gln Glu Ser Pro Lys Ala Thr Gly Val Phe Thr
                1045                1050                1055
Thr Leu Gln Pro Gly Ser Ser Ile Pro Pro Tyr Asn Thr Glu Val Thr
            1060                1065                1070
Glu Thr Thr Ile Val Ile Thr Trp Thr Pro Ala Pro Arg Ile Gly Phe
        1075                1080                1085
Lys Leu Gly Val Arg Pro Ser Gln Gly Gly Glu Ala Pro Arg Glu Val
    1090                1095                1100
Thr Ser Asp Ser Gly Ser Ile Val Val Ser Gly Leu Thr Pro Gly Val
1105                1110                1115                1120
Glu Tyr Val Tyr Thr Ile Gln Val Leu Arg Asp Gly Gln Glu Arg Asp
                1125                1130                1135
Ala Pro Ile Val Asn Lys Val Val Thr Pro Leu Ser Pro Pro Thr Asn
            1140                1145                1150
Leu His Leu Glu Ala Asn Pro Asp Thr Gly Val Leu Thr Val Ser Trp
        1155                1160                1165
Glu Arg Ser Thr Thr Pro Asp Ile Thr Gly Tyr Arg Ile Thr Thr Thr
    1170                1175                1180
Pro Thr Asn Gly Gln Gln Gly Asn Ser Leu Glu Glu Val Val His Ala
1185                1190                1195                1200
```

Asp Gln Ser Ser Cys Thr Phe Asp Asn Leu Ser Pro Gly Leu Glu Tyr
1205 1210 1215
Asn Val Ser Val Tyr Thr Val Lys Asp Asp Lys Glu Ser Val Pro Ile
1220 1225 1230
Ser Asp Thr Ile Ile Pro Ala Val Pro Pro Pro Thr Asp Leu Arg Phe
1235 1240 1245
Thr Asn Ile Gly Pro Asp Thr Met Arg Val Thr Trp Ala Pro Pro Pro
1250 1255 1260
Ser Ile Asp Leu Thr Asn Phe Leu Val Arg Tyr Ser Pro Val Lys Asn
1265 1270 1275 1280
Glu Glu Asp Val Ala Glu Leu Ser Ile Ser Pro Ser Asp Asn Ala Val
1285 1290 1295
Val Leu Thr Asn Leu Leu Pro Gly Thr Glu Tyr Val Val Ser Val Ser
1300 1305 1310
Ser Val Tyr Glu Gln His Glu Ser Thr Pro Leu Arg Gly Arg Gln Lys
1315 1320 1325
Thr Gly Leu Asp Ser Pro Thr Gly Ile Asp Phe Ser Asp Ile Thr Ala
1330 1335 1340
Asn Ser Phe Thr Val His Trp Ile Ala Pro Arg Ala Thr Ile Thr Gly
1345 1350 1355 1360
Tyr Arg Ile Arg His His Pro Glu His Phe Ser Gly Arg Pro Arg Glu
1365 1370 1375
Asp Arg Val Pro His Ser Arg Asn Ser Ile Thr Leu Thr Asn Leu Thr
1380 1385 1390
Pro Gly Thr Glu Tyr Val Val Ser Ile Val Ala Leu Asn Gly Arg Glu
1395 1400 1405
Glu Ser Pro Leu Leu Ile Gly Gln Gln Ser Thr Val Ser Asp Val Pro
1410 1415 1420
Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr Ser Leu Leu Ile Ser
1425 1430 1435 1440
Trp Asp Ala Pro Ala Val Thr Val Arg Tyr Tyr Arg Ile Thr Tyr Gly
1445 1450 1455
Glu Thr Gly Gly Asn Ser Pro Val Gln Glu Phe Thr Val Pro Gly Ser
1460 1465 1470
Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro Gly Val Asp Tyr Thr
1475 1480 1485
Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Asp Ser Pro Ala Ser Ser
1490 1495 1500
Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile Asp Lys Pro Ser Gln
1505 1510 1515 1520
Met Gln Val Thr Asp Val Gln Asp Asn Ser Ile Ser Val Lys Trp Leu
1525 1530 1535
Pro Ser Ser Ser Pro Val Thr Gly Tyr Arg Val Thr Thr Thr Pro Lys
1540 1545 1550
Asn Gly Pro Gly Pro Thr Lys Thr Lys Thr Ala Gly Pro Asp Gln Thr
1555 1560 1565
Glu Met Thr Ile Glu Gly Leu Gln Pro Thr Val Glu Tyr Val Val Ser
1570 1575 1580
Val Tyr Ala Gln Asn Pro Ser Gly Glu Ser Gln Pro Leu Val Gln Thr
1585 1590 1595 1600
Ala Val Thr Asn Ile Asp Arg Pro Lys Gly Leu Ala Phe Thr Asp Val
1605 1610 1615
Asp Val Asp Ser Ile Lys Ile Ala Trp Glu Ser Pro Gln Gly Gln Val
1620 1625 1630
Ser Arg Tyr Arg Val Thr Tyr Ser Ser Pro Glu Asp Gly Ile His Glu
1635 1640 1645
Leu Phe Pro Ala Pro Asp Gly Glu Glu Asp Thr Ala Glu Leu Gln Gly

1650 1655 1660
Leu Arg Pro Gly Ser Glu Tyr Thr Val Ser Val Val Ala Leu His Asp
1665 1670 1675 1680
Asp Met Glu Ser Gln Pro Leu Ile Gly Thr Gln Ser Thr Ala Ile Pro
1685 1690 1695
Ala Pro Thr Asp Leu Lys Phe Thr Gln Val Thr Pro Thr Ser Leu Ser
1700 1705 1710
Ala Gln Trp Thr Pro Pro Asn Val Gln Leu Thr Gly Tyr Arg Val Arg
1715 1720 1725
Val Thr Pro Lys Glu Lys Thr Gly Pro Met Lys Glu Ile Asn Leu Ala
1730 1735 1740
Pro Asp Ser Ser Ser Val Val Val Ser Gly Leu Met Val Ala Thr Lys
1745 1750 1755 1760
Tyr Glu Val Ser Val Tyr Ala Leu Lys Asp Thr Leu Thr Ser Arg Pro
1765 1770 1775
Ala Gln Gly Val Val Thr Thr Leu Glu Asn Val Ser Pro Pro Arg Arg
1780 1785 1790
Ala Arg Val Thr Asp Ala Thr Glu Thr Thr Ile Thr Ile Ser Trp Arg
1795 1800 1805
Thr Lys Thr Glu Thr Ile Thr Gly Phe Gln Val Asp Ala Val Pro Ala
1810 1815 1820
Asn Gly Gln Thr Pro Ile Gln Arg Thr Ile Lys Pro Asp Val Arg Ser
1825 1830 1835 1840
Tyr Thr Ile Thr Gly Leu Gln Pro Gly Thr Asp Tyr Lys Ile Tyr Leu
1845 1850 1855
Tyr Thr Leu Asn Asp Asn Ala Arg Ser Ser Pro Val Val Ile Asp Ala
1860 1865 1870
Ser Thr Ala Ile Asp Ala Pro Ser Asn Leu Arg Phe Leu Ala Thr Thr
1875 1880 1885
Pro Asn Ser Leu Leu Val Ser Trp Gln Pro Pro Arg Ala Arg Ile Thr
1890 1895 1900
Gly Tyr Ile Ile Lys Tyr Glu Lys Pro Gly Ser Pro Pro Arg Glu Val
1905 1910 1915 1920
Val Pro Arg Pro Arg Pro Gly Val Thr Glu Ala Thr Ile Thr Gly Leu
1925 1930 1935
Glu Pro Gly Thr Glu Tyr Thr Ile Tyr Val Ile Ala Leu Lys Asn Asn
1940 1945 1950
Gln Lys Ser Glu Pro Leu Ile Gly Arg Lys Lys Thr Asp Glu Leu Pro
1955 1960 1965
Gln Leu Val Thr Leu Pro His Pro Asn Leu His Gly Pro Glu Ile Leu
1970 1975 1980
Asp Val Pro Ser Thr Val Gln Lys Thr Pro Phe Val Thr His Pro Gly
1985 1990 1995 2000
Tyr Asp Thr Gly Asn Gly Ile Gln Leu Pro Gly Thr Ser Gly Gln Gln
2005 2010 2015
Pro Ser Val Gly Gln Gln Met Ile Phe Glu Glu His Gly Phe Arg Arg
2020 2025 2030
Thr Thr Pro Pro Thr Thr Ala Thr Pro Ile Arg His Arg Pro Arg Pro
2035 2040 2045
Tyr Pro Pro Asn Val Gly Gln Glu Ala Leu Ser Gln Thr Thr Ile Ser
2050 2055 2060
Trp Ala Pro Phe Gln Asp Thr Ser Glu Tyr Ile Ile Ser Cys His Pro
2065 2070 2075 2080
Val Gly Thr Asp Glu Glu Pro Leu Gln Phe Arg Val Pro Gly Thr Ser
2085 2090 2095
Thr Ser Ala Thr Leu Thr Gly Leu Thr Arg Gly Ala Thr Tyr Asn Ile
2100 2105 2110

Ile Val Glu Ala Leu Lys Asp Gln Gln Arg His Lys Val Arg Glu Glu
2115 2120 2125
Val Val Thr Val Gly Asn Ser Val Asn Glu Gly Leu Asn Gln Pro Thr
2130 2135 2140
Asp Asp Ser Cys Phe Asp Pro Tyr Thr Val Ser His Tyr Ala Val Gly
2145 2150 2155 2160
Asp Glu Trp Glu Arg Met Ser Glu Ser Gly Phe Lys Leu Leu Cys Gln
2165 2170 2175
Cys Leu Gly Phe Gly Ser Gly His Phe Arg Cys Asp Ser Ser Arg Trp
2180 2185 2190
Cys His Asp Asn Gly Val Asn Tyr Lys Ile Gly Glu Lys Trp Asp Arg
2195 2200 2205
Gln Gly Glu Asn Gly Gln Met Met Ser Cys Thr Cys Leu Gly Asn Gly
2210 2215 2220
Lys Gly Glu Phe Lys Cys Asp Pro His Glu Ala Thr Cys Tyr Asp Asp
2225 2230 2235 2240
Gly Lys Thr Tyr His Val Gly Glu Gln Trp Gln Lys Glu Tyr Leu Gly
2245 2250 2255
Ala Ile Cys Ser Cys Thr Cys Phe Gly Gly Gln Arg Gly Trp Arg Cys
2260 2265 2270
Asp Asn Cys Arg Arg Pro Gly Gly Glu Pro Ser Pro Glu Gly Thr Thr
2275 2280 2285
Gly Gln Ser Tyr Asn Gln Tyr Ser Gln Arg Tyr His Gln Arg Thr Asn
2290 2295 2300
Thr Asn Val Asn Cys Pro Ile Glu Cys Phe Met Pro Leu Asp Val Gln
2305 2310 2315 2320
Ala Asp Arg Glu Asp Ser Arg Glu
2325

<210> 99
<211> 188
<212> PRT
<213> Homo sapiens

<400> 99
His Gln Thr His Lys Glu Gly Gly Ser Thr His Ala Ser Ala Asp Ala
1 5 10 15
Trp Glu Ile Ile Glu Leu Glu Thr Glu Ile Glu Lys Phe Lys Ala Glu
20 25 30
Asn Ala Ser Leu Ala Lys Leu Arg Ile Glu Arg Glu Ser Ala Leu Glu
35 40 45
Lys Leu Arg Lys Glu Ile Ala Asp Phe Glu Gln Gln Lys Ala Lys Glu
50 55 60
Leu Ala Arg Ile Glu Glu Phe Lys Lys Glu Glu Met Arg Lys Leu Gln
65 70 75 80
Lys Glu Arg Lys Val Phe Glu Lys Tyr Thr Thr Ala Ala Arg Thr Phe
85 90 95
Pro Asp Lys Lys Glu Arg Glu Glu Ile Gln Thr Leu Lys Gln Gln Ile
100 105 110
Ala Asp Leu Arg Glu Asp Leu Lys Arg Lys Glu Thr Lys Trp Ser Ser
115 120 125
Thr His Ser Arg Leu Arg Ser Gln Ile Gln Met Leu Val Arg Glu Asn
130 135 140
Thr Asp Leu Arg Glu Glu Ile Lys Val Met Glu Arg Phe Arg Leu Asp
145 150 155 160
Ala Trp Lys Arg Ala Glu Ala Ile Glu Ser Ser Leu Glu Val Glu Lys
165 170 175

Lys Asp Lys Leu Ala Asn Thr Ser Val Arg Phe Gln
180 185

<210> 100
<211> 284
<212> PRT
<213> Homo sapiens

<400> 100
Met Glu Pro Gly Asn Tyr Ala Thr Leu Asp Gly Ala Lys Asp Ile Glu
1 5 10 15
Gly Leu Leu Gly Ala Gly Gly Gly Arg Asn Leu Val Ala His Ser Pro
20 25 30
Leu Thr Ser His Pro Ala Ala Pro Thr Leu Met Pro Ala Val Asn Tyr
35 40 45
Ala Pro Leu Asp Leu Pro Gly Ser Ala Glu Pro Pro Lys Gln Cys His
50 55 60
Pro Cys Pro Gly Val Pro Gln Gly Thr Ser Pro Ala Pro Val Pro Tyr
65 70 75 80
Gly Tyr Phe Gly Gly Gly Tyr Tyr Ser Cys Arg Val Ser Arg Ser Ser
85 90 95
Leu Lys Pro Cys Ala Gln Ala Ala Thr Leu Ala Ala Tyr Pro Ala Glu
100 105 110
Thr Pro Thr Ala Gly Glu Glu Tyr Pro Ser Arg Pro Thr Glu Phe Ala
115 120 125
Phe Tyr Pro Gly Tyr Pro Gly Thr Tyr His Ala Met Ala Ser Tyr Leu
130 135 140
Asp Val Ser Val Val Gln Thr Leu Gly Ala Pro Gly Glu Pro Arg His
145 150 155 160
Asp Ser Leu Leu Pro Val Asp Ser Tyr Gln Ser Trp Ala Leu Ala Gly
165 170 175
Gly Trp Asn Ser Gln Met Cys Cys Gln Gly Glu Gln Asn Pro Pro Gly
180 185 190
Pro Phe Trp Lys Ala Ala Phe Ala Asp Ser Ser Gly Gln His Pro Pro
195 200 205
Asp Ala Cys Ala Phe Arg Arg Gly Arg Lys Lys Arg Ile Pro Tyr Ser
210 215 220
Lys Gly Gln Leu Arg Glu Leu Glu Arg Glu Tyr Ala Ala Asn Lys Phe
225 230 235 240
Ile Thr Lys Asp Lys Arg Arg Lys Ile Ser Ala Ala Thr Ser Leu Ser
245 250 255
Glu Arg Gln Ile Thr Ile Trp Phe Gln Asn Arg Arg Val Lys Glu Lys
260 265 270
Lys Val Leu Ala Lys Val Lys Asn Ser Ala Thr Pro
275 280

<210> 101
<211> 676
<212> PRT
<213> Homo sapiens

<400> 101
Met Asp Lys Tyr Asp Asp Leu Gly Leu Glu Ala Ser Lys Phe Ile Glu
1 5 10 15
Asp Leu Asn Met Tyr Glu Ala Ser Lys Asp Gly Leu Phe Arg Val Asp
20 25 30
Lys Gly Ala Gly Asn Asn Pro Glu Phe Glu Glu Thr Arg Arg Val Phe

```
            35                  40                  45
Ala Thr Lys Met Ala Lys Ile His Leu Gln Gln Gln Gln Gln Gln Leu
        50                  55                  60
Leu Gln Glu Glu Thr Leu Pro Arg Gly Ser Arg Gly Pro Val Asn Gly
65                  70                  75                  80
Gly Gly Arg Leu Gly Pro Gln Ala Arg Trp Glu Val Val Gly Ser Lys
                85                  90                  95
Leu Thr Val Asp Gly Ala Ala Lys Pro Pro Leu Ala Ala Ser Thr Gly
            100                 105                 110
Ala Pro Gly Ala Val Thr Thr Leu Ala Ala Gly Gln Pro Pro Tyr Pro
        115                 120                 125
Pro Gln Glu Gln Arg Ser Arg Pro Tyr Leu His Gly Thr Arg His Gly
    130                 135                 140
Ser Gln Asp Cys Gly Ser Arg Glu Ser Leu Ala Thr Ser Glu Met Ser
145                 150                 155                 160
Ala Phe His Gln Pro Gly Pro Cys Glu Asp Pro Ser Cys Leu Thr His
                165                 170                 175
Gly Asp Tyr Tyr Asp Asn Leu Ser Leu Ala Ser Pro Lys Trp Gly Asp
            180                 185                 190
Lys Pro Gly Val Ser Pro Ser Ile Gly Leu Ser Val Gly Ser Gly Trp
        195                 200                 205
Pro Ser Ser Pro Gly Ser Asp Pro Pro Leu Pro Lys Pro Cys Gly Asp
    210                 215                 220
His Pro Leu Asn His Arg Gln Leu Ser Leu Ser Ser Ser Arg Ser Ser
225                 230                 235                 240
Glu Gly Ser Leu Gly Gly Gln Asn Ser Gly Ile Gly Gly Arg Ser Ser
                245                 250                 255
Glu Lys Pro Thr Gly Leu Trp Ser Thr Ala Ser Ser Gln Arg Val Ser
            260                 265                 270
Pro Gly Leu Pro Ser Pro Asn Leu Glu Asn Gly Ala Pro Ala Val Gly
        275                 280                 285
Pro Val Gln Pro Arg Thr Pro Ser Val Ser Ala Pro Leu Ala Leu Ser
    290                 295                 300
Cys Pro Arg Gln Gly Gly Leu Pro Arg Ser Asn Ser Gly Leu Gly Gly
305                 310                 315                 320
Glu Val Ser Gly Val Met Ser Lys Pro Asn Val Asp Pro Gln Pro Trp
                325                 330                 335
Phe Gln Asp Gly Pro Lys Ser Tyr Leu Ser Ser Ser Ala Pro Ser Ser
            340                 345                 350
Ser Pro Ala Gly Leu Asp Gly Ser Gln Gln Gly Ala Val Pro Gly Leu
        355                 360                 365
Gly Pro Lys Pro Gly Cys Thr Asp Leu Gly Thr Gly Pro Lys Leu Ser
    370                 375                 380
Pro Thr Ser Leu Val His Pro Val Met Ser Thr Leu Pro Glu Leu Ser
385                 390                 395                 400
Cys Lys Glu Gly Pro Leu Gly Trp Ser Ser Asp Gly Ser Leu Gly Ser
                405                 410                 415
Val Leu Leu Asp Ser Pro Ser Ser Pro Arg Val Arg Leu Pro Cys Gln
            420                 425                 430
Pro Leu Val Pro Gly Pro Glu Leu Arg Pro Ser Ala Ala Glu Leu Lys
        435                 440                 445
Leu Glu Ala Leu Thr Gln Arg Leu Glu Arg Glu Met Asp Ala His Pro
    450                 455                 460
Lys Ala Asp Tyr Phe Gly Ala Cys Val Lys Cys Ser Lys Gly Val Phe
465                 470                 475                 480
Gly Ala Gly Gln Ala Cys Gln Ala Met Gly Asn Leu Tyr His Asp Thr
                485                 490                 495
```

```
Cys Phe Thr Cys Ala Ala Cys Ser Arg Lys Leu Arg Gly Lys Ala Phe
            500                 505                 510
Tyr Phe Val Asn Gly Lys Val Phe Cys Glu Glu Asp Phe Leu Tyr Ser
        515                 520                 525
Gly Phe Gln Gln Ser Ala Asp Arg Cys Phe Leu Cys Gly His Leu Ile
    530                 535                 540
Met Asp Met Ile Leu Gln Ala Leu Gly Lys Ser Tyr His Pro Gly Cys
545                 550                 555                 560
Phe Arg Cys Val Ile Cys Asn Glu Cys Leu Asp Gly Val Pro Phe Thr
                565                 570                 575
Val Asp Ser Glu Asn Lys Ile Tyr Cys Val Arg Asp Tyr His Lys Val
            580                 585                 590
Leu Ala Pro Lys Cys Ala Ala Cys Gly Leu Pro Ile Leu Pro Pro Glu
        595                 600                 605
Gly Ser Asp Glu Thr Ile Arg Val Val Ser Met Asp Arg Asp Tyr His
    610                 615                 620
Val Glu Cys Tyr His Cys Glu Asp Cys Gly Leu Glu Leu Asn Asp Glu
625                 630                 635                 640
Asp Gly His Arg Cys Tyr Pro Leu Glu Asp His Leu Phe Cys His Ser
                645                 650                 655
Cys His Val Lys Arg Leu Glu Lys Arg Pro Ser Ser Thr Ala Leu His
            660                 665                 670
Gln His His Phe
        675

<210> 102
<211> 296
<212> PRT
<213> Homo sapiens

<400> 102
Ser Thr Gly Ser Glu Phe Pro Leu Cys Thr Lys Ala Ser Pro Cys Ser
1               5                   10                  15
Ala Ala Arg Ala Gly Gly Arg Ala Leu Gly Trp Arg Leu Gln Gln Gln
            20                  25                  30
Arg Glu Thr Arg Gly Asn Pro Gly Asn Pro Gly Leu Gly Val Ala Ala
        35                  40                  45
Thr Met Thr Gly Ser Asn Met Ser Asp Ala Leu Ala Asn Ala Val Cys
    50                  55                  60
Gln Arg Cys Gln Ala Arg Phe Ser Pro Ala Glu Arg Ile Val Asn Ser
65                  70                  75                  80
Asn Gly Glu Leu Tyr His Glu His Cys Phe Val Cys Ala Gln Cys Phe
                85                  90                  95
Arg Pro Phe Pro Glu Gly Leu Phe Tyr Glu Phe Glu Gly Arg Lys Tyr
            100                 105                 110
Cys Glu His Asp Phe Gln Met Leu Phe Ala Pro Cys Cys Gly Ser Cys
        115                 120                 125
Gly Glu Phe Ile Ile Gly Arg Val Ile Lys Ala Met Asn Asn Asn Trp
    130                 135                 140
His Pro Gly Cys Phe Arg Cys Glu Leu Cys Asp Val Glu Leu Ala Asp
145                 150                 155                 160
Leu Gly Phe Val Lys Asn Ala Gly Arg His Leu Cys Arg Pro Cys His
                165                 170                 175
Asn Arg Glu Lys Ala Lys Gly Leu Gly Lys Tyr Ile Cys Gln Arg Cys
            180                 185                 190
His Leu Val Ile Asp Glu Gln Pro Leu Met Phe Arg Ser Asp Ala Tyr
        195                 200                 205
```

```
His Pro Asp His Phe Asn Cys Thr His Cys Gly Lys Glu Leu Thr Ala
    210                 215                 220
Glu Ala Arg Glu Leu Lys Gly Glu Leu Tyr Cys Leu Pro Cys His Asp
225                 230                 235                 240
Lys Met Gly Val Pro Ile Cys Gly Ala Cys Arg Arg Pro Ile Glu Gly
                245                 250                 255
Arg Val Val Asn Ala Leu Gly Lys Gln Trp His Val Glu His Phe Val
            260                 265                 270
Cys Ala Lys Cys Glu Lys Pro Phe Leu Gly His Arg His Tyr Glu Lys
        275                 280                 285
Lys Gly Leu Ala Tyr Cys Glu Leu
    290                 295

&lt;210&gt; 103
&lt;211&gt; 500
&lt;212&gt; PRT
&lt;213&gt; Homo sapiens

&lt;400&gt; 103
Met Gly Ile Gly Leu Ser Ala Gln Gly Val Asn Met Asn Arg Leu Pro
1               5                   10                  15
Gly Trp Asp Lys His Ser Tyr Gly Tyr His Gly Asp Asp Gly His Ser
            20                  25                  30
Phe Cys Ser Ser Gly Thr Gly Gln Pro Tyr Gly Pro Thr Phe Thr Thr
        35                  40                  45
Gly Asp Val Ile Gly Cys Cys Val Asn Leu Ile Asn Asn Thr Cys Phe
    50                  55                  60
Tyr Thr Lys Asn Gly His Ser Leu Gly Ile Ala Phe Thr Asp Leu Pro
65                  70                  75                  80
Pro Asn Leu Tyr Pro Thr Val Gly Leu Gln Thr Pro Gly Glu Val Val
                85                  90                  95
Asp Ala Asn Phe Gly Gln His Pro Phe Val Phe Asp Ile Glu Asp Tyr
            100                 105                 110
Met Arg Glu Trp Arg Thr Lys Ile Gln Ala Gln Ile Asp Arg Phe Pro
        115                 120                 125
Ile Gly Asp Arg Glu Gly Glu Trp Gln Thr Met Ile Gln Lys Met Val
    130                 135                 140
Ser Ser Tyr Leu Val His His Gly Tyr Cys Ala Thr Ala Glu Ala Phe
145                 150                 155                 160
Ala Arg Ser Thr Asp Gln Thr Val Leu Glu Glu Leu Ala Ser Ile Lys
                165                 170                 175
Asn Arg Gln Arg Ile Gln Lys Leu Val Leu Ala Gly Arg Met Gly Glu
            180                 185                 190
Ala Ile Glu Thr Thr Gln Gln Leu Tyr Pro Ser Leu Leu Glu Arg Asn
        195                 200                 205
Pro Asn Leu Leu Phe Thr Leu Lys Val Arg Gln Phe Ile Glu Met Val
    210                 215                 220
Asn Gly Thr Asp Ser Glu Val Arg Cys Leu Gly Gly Arg Ser Pro Lys
225                 230                 235                 240
Ser Gln Asp Ser Tyr Pro Val Ser Pro Arg Pro Phe Ser Ser Pro Ser
                245                 250                 255
Met Ser Pro Ser His Gly Met Asn Ile His Asn Leu Ala Ser Gly Lys
            260                 265                 270
Gly Ser Thr Ala His Phe Ser Gly Phe Glu Ser Cys Ser Asn Gly Val
        275                 280                 285
Ile Ser Asn Lys Ala His Gln Ser Tyr Cys His Ser Asn Lys His Gln
    290                 295                 300
```

Ser Ser Asn Leu Asn Val Pro Glu Leu Asn Ser Ile Asn Met Ser Arg
305 310 315 320
Ser Gln Gln Val Asn Asn Phe Thr Ser Asn Asp Val Asp Met Glu Thr
325 330 335
Asp His Tyr Ser Asn Gly Val Gly Glu Thr Ser Ser Asn Gly Phe Leu
340 345 350
Asn Gly Ser Ser Lys His Asp His Glu Met Glu Asp Cys Asp Thr Glu
355 360 365
Met Glu Val Asp Ser Ser Gln Leu Arg Arg Gln Leu Cys Gly Gly Ser
370 375 380
Gln Ala Ala Ile Glu Arg Met Ile His Phe Gly Arg Glu Leu Gln Ala
385 390 395 400
Met Ser Glu Gln Leu Arg Arg Asp Cys Gly Lys Asn Thr Ala Asn Lys
405 410 415
Lys Met Leu Lys Asp Ala Phe Ser Leu Leu Ala Tyr Ser Asp Pro Trp
420 425 430
Asn Ser Pro Val Gly Asn Gln Leu Asp Pro Ile Gln Arg Glu Pro Val
435 440 445
Cys Ser Ala Leu Asn Ser Ala Ile Leu Glu Thr His Asn Leu Pro Lys
450 455 460
Gln Pro Pro Leu Ala Leu Ala Met Gly Gln Ala Thr Gln Cys Leu Gly
465 470 475 480
Leu Met Ala Arg Ser Gly Ile Gly Ser Cys Ala Phe Ala Thr Val Glu
485 490 495
Asp Tyr Leu His
500

<210> 104
<211> 387
<212> PRT
<213> Homo sapiens

<400> 104
Met Ala Thr Ser Gly Val Leu Pro Gly Gly Gly Phe Val Ala Ser Ala
1 5 10 15
Ala Ala Val Ala Gly Pro Glu Met Gln Thr Gly Arg Asn Asn Phe Val
20 25 30
Ile Arg Arg Asn Pro Ala Asp Pro Gln Arg Ile Pro Ser Asn Pro Ser
35 40 45
His Arg Ile Gln Cys Ala Ala Gly Tyr Glu Gln Ser Glu His Asn Val
50 55 60
Cys Gln Asp Ile Asp Glu Cys Thr Ala Gly Thr His Asn Cys Arg Ala
65 70 75 80
Asp Gln Val Cys Ile Asn Leu Arg Gly Ser Phe Ala Cys Gln Cys Pro
85 90 95
Pro Gly Tyr Gln Lys Arg Gly Glu Gln Cys Val Asp Ile Asp Glu Cys
100 105 110
Thr Ile Pro Pro Tyr Cys His Gln Arg Cys Val Asn Thr Pro Gly Ser
115 120 125
Phe Tyr Cys Gln Cys Ser Pro Gly Phe Gln Leu Ala Ala Asn Asn Tyr
130 135 140
Thr Cys Val Asp Ile Asn Glu Cys Asp Ala Ser Asn Gln Cys Ala Gln
145 150 155 160
Gln Cys Tyr Asn Ile Leu Gly Ser Phe Ile Cys Gln Cys Asn Gln Gly
165 170 175
Tyr Glu Leu Ser Ser Asp Arg Leu Asn Cys Glu Asp Ile Asp Glu Cys
180 185 190

Arg Thr Ser Ser Tyr Leu Cys Gln Tyr Gln Cys Val Asn Glu Pro Gly
195 200 205
Lys Phe Ser Cys Met Cys Pro Gln Gly Tyr Gln Val Val Arg Ser Arg
210 215 220
Thr Cys Gln Asp Ile Asn Glu Cys Glu Thr Thr Asn Glu Cys Arg Glu
225 230 235 240
Asp Glu Met Cys Trp Asn Tyr His Gly Gly Phe Arg Cys Tyr Pro Arg
245 250 255
Asn Pro Cys Gln Asp Pro Tyr Ile Leu Thr Pro Glu Asn Arg Cys Val
260 265 270
Cys Pro Val Ser Asn Ala Met Cys Arg Glu Leu Pro Gln Ser Ile Val
275 280 285
Tyr Lys Tyr Met Ser Ile Arg Ser Asp Arg Ser Val Pro Ser Asp Ile
290 295 300
Phe Gln Ile Gln Ala Thr Thr Ile Tyr Ala Asn Thr Ile Asn Thr Phe
305 310 315 320
Arg Ile Lys Ser Gly Asn Glu Asn Gly Glu Phe Tyr Leu Arg Gln Thr
325 330 335
Ser Pro Val Ser Ala Met Leu Val Leu Val Lys Ser Leu Ser Gly Pro
340 345 350
Arg Glu His Ile Val Asp Leu Glu Met Leu Thr Val Ser Ser Ile Gly
355 360 365
Thr Phe Arg Thr Ser Ser Val Leu Arg Leu Thr Ile Ile Val Gly Pro
370 375 380
Phe Ser Phe
385

<210> 105
<211> 531
<212> PRT
<213> Homo sapiens

<400> 105
Met Ser Lys Pro His Ser Glu Ala Gly Thr Ala Phe Ile Gln Thr Gln
1 5 10 15
Gln Leu His Ala Ala Met Ala Asp Thr Phe Leu Glu His Met Cys Arg
20 25 30
Leu Asp Ile Asp Ser Pro Pro Ile Thr Ala Arg Asn Thr Gly Ile Ile
35 40 45
Cys Thr Ile Gly Pro Ala Ser Arg Ser Val Glu Thr Leu Lys Glu Met
50 55 60
Ile Lys Ser Gly Met Asn Val Ala Arg Leu Asn Phe Ser His Gly Thr
65 70 75 80
His Glu Tyr His Ala Glu Thr Ile Lys Asn Val Arg Thr Ala Thr Glu
85 90 95
Ser Phe Ala Ser Asp Pro Tyr Leu Tyr Arg Pro Val Ala Val Ala Leu
100 105 110
Asp Thr Lys Gly Pro Glu Ile Arg Thr Gly Leu Ile Lys Gly Ser Gly
115 120 125
Thr Ala Glu Leu Glu Leu Lys Lys Gly Ala Thr Leu Lys Ile Thr Leu
130 135 140
Asp Asn Ala Tyr Met Glu Lys Cys Asp Glu Asn Ile Leu Trp Leu Asp
145 150 155 160
Tyr Lys Asn Ile Cys Lys Val Val Glu Val Gly Ser Lys Ile Tyr Val
165 170 175
Asp Asp Gly Leu Ile Ser Leu Gln Val Lys Gln Lys Gly Ala Asp Phe
180 185 190

Leu Val Thr Glu Val Glu Asn Gly Gly Ser Leu Gly Ser Lys Lys Gly
        195   200   205
Val Asn Leu Pro Gly Ala Ala Val Asp Leu Pro Ala Val Ser Glu Lys
    210   215   220
Asp Ile Gln Asp Leu Lys Phe Gly Val Glu Gln Asp Val Asp Met Val
225   230   235   240
Phe Ala Ser Phe Ile Arg Lys Ala Ser Asp Val His Glu Val Arg Lys
                245   250   255
Val Leu Gly Glu Lys Gly Lys Asn Ile Lys Ile Ile Ser Lys Ile Glu
            260   265   270
Asn His Glu Gly Val Arg Arg Phe Asp Glu Ile Leu Glu Ala Ser Asp
        275   280   285
Gly Ile Met Val Ala Arg Gly Asp Leu Gly Ile Glu Ile Pro Ala Glu
    290   295   300
Lys Val Phe Leu Ala Gln Lys Met Met Ile Gly Arg Cys Asn Arg Ala
305   310   315   320
Gly Lys Pro Val Ile Cys Ala Thr Gln Met Leu Glu Ser Met Ile Lys
                325   330   335
Lys Pro Arg Pro Thr Arg Ala Glu Gly Ser Asp Val Ala Asn Ala Val
            340   345   350
Leu Asp Gly Ala Asp Cys Ile Met Leu Ser Gly Glu Thr Ala Lys Gly
        355   360   365
Asp Tyr Pro Leu Glu Ala Val Arg Met Gln His Leu Ile Ala Arg Glu
    370   375   380
Ala Glu Ala Ala Ile Tyr His Leu Gln Leu Phe Glu Glu Leu Arg Arg
385   390   395   400
Leu Ala Pro Ile Thr Ser Asp Pro Thr Glu Ala Thr Ala Val Gly Ala
                405   410   415
Val Glu Ala Ser Phe Lys Cys Cys Ser Gly Ala Ile Ile Val Leu Thr
            420   425   430
Lys Ser Gly Arg Ser Ala His Gln Val Ala Arg Tyr Arg Pro Arg Ala
        435   440   445
Pro Ile Ile Ala Val Thr Arg Asn Pro Gln Thr Ala Arg Gln Ala His
    450   455   460
Leu Tyr Arg Gly Ile Phe Pro Val Leu Cys Lys Asp Pro Val Gln Glu
465   470   475   480
Ala Trp Ala Glu Asp Val Asp Leu Arg Val Asn Phe Ala Met Asn Val
                485   490   495
Gly Lys Ala Arg Gly Phe Phe Lys Lys Gly Asp Val Val Ile Val Leu
            500   505   510
Thr Gly Trp Arg Pro Gly Ser Gly Phe Thr Asn Thr Met Arg Val Val
        515   520   525
Pro Val Pro
    530

<210> 106
<211> 480
<212> PRT
<213> Homo sapiens

<400> 106
Met Ala Ala Arg Cys Ser Thr Arg Trp Leu Leu Val Val Val Gly Thr
1   5   10   15
Pro Arg Leu Pro Ala Ile Ser Gly Arg Gly Ala Arg Pro Pro Arg Glu
            20   25   30
Gly Val Val Gly Ala Trp Leu Ser Arg Lys Leu Ser Val Pro Ala Phe
        35   40   45

```
Ala Ser Ser Leu Thr Ser Cys Gly Pro Arg Ala Leu Leu Thr Leu Arg
    50                  55                  60
Pro Gly Val Ser Leu Thr Gly Thr Lys His Asn Pro Phe Ile Cys Thr
65                  70                  75                  80
Ala Ser Phe His Thr Ser Ala Pro Leu Ala Lys Glu Asp Tyr Tyr Gln
                85                  90                  95
Ile Leu Gly Val Pro Arg Asn Ala Ser Gln Lys Glu Ile Lys Lys Ala
            100                 105                 110
Tyr Tyr Gln Leu Ala Lys Lys Tyr His Pro Asp Thr Asn Lys Asp Asp
        115                 120                 125
Pro Lys Ala Lys Glu Lys Phe Ser Gln Leu Ala Glu Ala Tyr Glu Val
    130                 135                 140
Leu Ser Asp Glu Val Lys Arg Lys Gln Tyr Asp Ala Tyr Gly Ser Ala
145                 150                 155                 160
Gly Phe Asp Pro Gly Ala Ser Gly Ser Gln His Ser Tyr Trp Lys Gly
                165                 170                 175
Gly Pro Thr Val Asp Pro Glu Glu Leu Phe Arg Lys Ile Phe Gly Glu
            180                 185                 190
Phe Ser Ser Ser Ser Phe Gly Asp Phe Gln Thr Val Phe Asp Gln Pro
        195                 200                 205
Gln Glu Tyr Phe Met Glu Leu Thr Phe Asn Gln Ala Ala Lys Gly Val
    210                 215                 220
Asn Lys Glu Phe Thr Val Asn Ile Met Asp Thr Cys Glu Arg Cys Asn
225                 230                 235                 240
Gly Lys Gly Asn Glu Pro Gly Thr Lys Val Gln His Cys His Tyr Cys
                245                 250                 255
Gly Gly Ser Gly Met Glu Thr Ile Asn Thr Gly Pro Phe Val Met Arg
            260                 265                 270
Ser Thr Cys Arg Arg Cys Gly Gly Arg Gly Ser Ile Ile Ile Ser Pro
        275                 280                 285
Cys Val Val Cys Arg Gly Ala Gly Gln Ala Lys Gln Lys Lys Arg Val
    290                 295                 300
Met Ile Pro Val Pro Ala Gly Val Glu Asp Gly Gln Thr Val Arg Met
305                 310                 315                 320
Pro Val Gly Lys Arg Glu Ile Phe Ile Thr Phe Arg Val Gln Lys Ser
                325                 330                 335
Pro Val Phe Arg Arg Asp Gly Ala Asp Ile His Ser Asp Leu Phe Ile
            340                 345                 350
Ser Ile Ala Gln Ala Leu Leu Gly Gly Thr Ala Arg Ala Gln Gly Leu
        355                 360                 365
Tyr Glu Thr Ile Asn Val Thr Ile Pro Pro Gly Thr Gln Thr Asp Gln
    370                 375                 380
Lys Ile Arg Met Gly Gly Lys Gly Ile Pro Arg Ile Asn Ser Tyr Gly
385                 390                 395                 400
Tyr Gly Asp His Tyr Ile His Ile Lys Ile Arg Val Pro Lys Arg Leu
                405                 410                 415
Thr Ser Arg Gln Gln Ser Leu Ile Leu Ser Tyr Ala Glu Asp Glu Thr
            420                 425                 430
Asp Val Glu Gly Thr Val Asn Gly Val Thr Leu Thr Ser Ser Gly Gly
        435                 440                 445
Ser Thr Met Asp Ser Ser Ala Gly Ser Lys Ala Arg Arg Glu Ala Gly
    450                 455                 460
Glu Asp Glu Glu Gly Phe Leu Ser Lys Leu Lys Lys Met Phe Thr Ser
465                 470                 475                 480
```

<210> 107
<211> 572

<212> PRT
<213> Homo sapiens

<400> 107
Met Ala Ala Pro Arg Pro Ser Pro Ala Ile Ser Val Ser Val Ser Ala
1   5   10   15
Pro Ala Phe Tyr Ala Pro Gln Lys Lys Phe Gly Pro Val Val Ala Pro
            20   25   30
Lys Pro Lys Val Asn Pro Phe Arg Pro Gly Asp Ser Glu Pro Pro Pro
        35   40   45
Ala Pro Gly Ala Gln Arg Ala Gln Met Gly Arg Val Gly Glu Ile Pro
    50   55   60
Pro Pro Pro Pro Glu Asp Phe Pro Leu Pro Pro Pro Pro Leu Ala Gly
65   70   75   80
Asp Gly Asp Asp Ala Glu Gly Ala Leu Gly Gly Ala Phe Pro Pro Pro
                85   90   95
Pro Pro Pro Ile Glu Glu Ser Phe Pro Pro Ala Pro Leu Glu Glu Glu
            100   105   110
Ile Phe Pro Ser Pro Pro Pro Pro Pro Glu Glu Glu Gly Gly Pro Glu
        115   120   125
Ala Pro Ile Pro Pro Pro Pro Gln Pro Arg Glu Lys Val Ser Ser Ile
    130   135   140
Asp Leu Glu Ile Asp Ser Leu Ser Ser Leu Leu Asp Asp Met Thr Lys
145   150   155   160
Asn Asp Pro Phe Lys Ala Arg Val Ser Ser Gly Tyr Val Pro Pro Pro
                165   170   175
Val Ala Thr Pro Phe Ser Ser Lys Ser Ser Thr Lys Pro Ala Ala Gly
            180   185   190
Gly Thr Ala Pro Leu Pro Pro Trp Lys Ser Pro Ser Ser Ser Gln Pro
        195   200   205
Leu Pro Gln Val Pro Ala Pro Ala Gln Ser Gln Thr Gln Phe His Val
    210   215   220
Gln Pro Gln Pro Gln Pro Lys Pro Gln Val Gln Leu His Val Gln Ser
225   230   235   240
Gln Thr Gln Pro Val Ser Leu Ala Asn Thr Gln Pro Arg Gly Pro Pro
                245   250   255
Ala Ser Ser Pro Ala Pro Ala Pro Lys Phe Ser Pro Val Thr Pro Lys
            260   265   270
Phe Thr Pro Val Ala Ser Lys Phe Ser Pro Gly Ala Pro Gly Gly Ser
        275   280   285
Gly Ser Gln Pro Asn Gln Lys Leu Gly His Pro Glu Ala Leu Ser Ala
    290   295   300
Gly Thr Gly Ser Pro Gln Pro Pro Ser Phe Thr Tyr Ala Gln Gln Arg
305   310   315   320
Glu Lys Pro Arg Val Gln Glu Lys Gln His Pro Val Pro Pro Pro Ala
                325   330   335
Gln Asn Gln Asn Gln Val Arg Ser Pro Gly Ala Pro Gly Pro Leu Thr
            340   345   350
Leu Lys Glu Val Glu Glu Leu Glu Gln Leu Thr Gln Gln Leu Met Gln
        355   360   365
Asp Met Glu His Pro Gln Arg Gln Asn Val Ala Val Asn Glu Leu Cys
    370   375   380
Gly Arg Cys His Gln Pro Leu Ala Arg Ala Gln Pro Ala Val Arg Ala
385   390   395   400
Leu Gly Gln Leu Phe His Ile Ala Cys Phe Thr Cys His Gln Cys Ala
                405   410   415
Gln Gln Leu Gln Gly Gln Gln Phe Tyr Ser Leu Glu Gly Ala Pro Tyr

```
            420                 425                 430
Cys Glu Gly Cys Tyr Thr Asp Thr Leu Glu Lys Cys Asn Thr Cys Gly
        435                 440                 445
Glu Pro Ile Thr Asp Arg Met Leu Arg Ala Thr Gly Lys Ala Tyr His
    450                 455                 460
Pro His Cys Phe Thr Cys Val Val Cys Ala Arg Pro Leu Glu Gly Thr
465                 470                 475                 480
Ser Phe Ile Val Asp Gln Ala Asn Arg Pro His Cys Val Pro Asp Tyr
                485                 490                 495
His Lys Gln Tyr Ala Pro Arg Cys Ser Val Cys Ser Glu Pro Ile Met
            500                 505                 510
Pro Glu Pro Gly Arg Asp Glu Thr Val Arg Val Val Ala Leu Asp Lys
        515                 520                 525
Asn Phe His Met Lys Cys Tyr Lys Cys Glu Asp Cys Gly Lys Pro Leu
    530                 535                 540
Ser Ile Glu Ala Asp Asp Asn Gly Cys Phe Pro Leu Asp Gly His Val
545                 550                 555                 560
Leu Cys Arg Lys Cys His Thr Ala Arg Ala Gln Thr
                565                 570

<210> 108
<211> 2861
<212> PRT
<213> Homo sapiens

<400> 108
Met Lys Ala Met Asp Val Leu Pro Ile Leu Lys Glu Lys Val Ala Tyr
1               5                   10                  15
Leu Ser Gly Gly Arg Asp Lys Arg Gly Gly Pro Ile Leu Thr Phe Pro
            20                  25                  30
Ala Arg Ser Asn His Asp Arg Ile Arg Gln Glu Asp Leu Arg Arg Leu
        35                  40                  45
Ile Ser Tyr Leu Ala Cys Ile Pro Ser Glu Glu Val Cys Lys Arg Gly
    50                  55                  60
Phe Thr Val Ile Val Asp Met Arg Gly Ser Lys Trp Asp Ser Ile Lys
65                  70                  75                  80
Pro Leu Leu Lys Ile Leu Gln Glu Ser Phe Pro Cys Cys Ile His Val
                85                  90                  95
Ala Leu Ile Ile Lys Pro Asp Asn Phe Trp Gln Lys Gln Arg Thr Asn
            100                 105                 110
Phe Gly Ser Ser Lys Phe Glu Phe Glu Thr Asn Met Val Ser Leu Glu
        115                 120                 125
Gly Leu Thr Lys Val Val Asp Pro Ser Gln Leu Thr Pro Glu Phe Asp
    130                 135                 140
Gly Cys Leu Glu Tyr Asn His Glu Glu Trp Ile Glu Ile Arg Val Ala
145                 150                 155                 160
Phe Glu Asp Tyr Ile Ser Asn Ala Thr His Met Leu Ser Arg Leu Glu
                165                 170                 175
Glu Leu Gln Asp Ile Leu Ala Lys Lys Glu Leu Pro Gln Asp Leu Glu
            180                 185                 190
Gly Ala Arg Asn Met Ile Glu Glu His Ser Gln Leu Lys Lys Lys Val
        195                 200                 205
Ile Lys Ala Pro Ile Glu Asp Leu Asp Leu Glu Gly Gln Lys Leu Leu
    210                 215                 220
Gln Arg Ile Gln Ser Ser Glu Ser Phe Pro Lys Lys Asn Ser Gly Ser
225                 230                 235                 240
Gly Asn Ala Asp Leu Gln Asn Leu Leu Pro Lys Val Ser Thr Met Leu
```

```
                245                 250                 255
Asp Arg Leu His Ser Thr Arg Gln His Leu His Gln Met Trp His Val
            260                 265                 270
Arg Lys Leu Lys Leu Asp Gln Cys Phe Gln Leu Arg Leu Phe Glu Gln
        275                 280                 285
Asp Ala Glu Lys Met Phe Asp Trp Ile Thr His Asn Lys Gly Leu Phe
    290                 295                 300
Leu Asn Ser Tyr Thr Glu Ile Gly Thr Ser His Pro His Ala Met Glu
305                 310                 315                 320
Leu Gln Thr Gln His Asn His Phe Ala Met Asn Cys Met Asn Val Tyr
                325                 330                 335
Val Asn Ile Asn Arg Ile Met Ser Val Ala Asn Arg Leu Val Glu Ser
            340                 345                 350
Gly His Tyr Ala Ser Gln Gln Ile Arg Gln Ile Ala Ser Gln Leu Glu
        355                 360                 365
Gln Glu Trp Lys Ala Phe Ala Ala Ala Leu Asp Glu Arg Ser Thr Leu
    370                 375                 380
Leu Asp Met Ser Ser Ile Phe His Gln Lys Ala Glu Lys Tyr Met Ser
385                 390                 395                 400
Asn Val Asp Ser Trp Cys Lys Ala Cys Gly Glu Val Asp Leu Pro Ser
                405                 410                 415
Glu Leu Gln Asp Leu Glu Asp Ala Ile His His His Gln Gly Ile Tyr
            420                 425                 430
Glu His Ile Thr Leu Ala Tyr Ser Glu Val Ser Gln Asp Gly Lys Ser
        435                 440                 445
Leu Leu Asp Lys Leu Gln Arg Pro Leu Thr Pro Gly Ser Ser Asp Ser
    450                 455                 460
Leu Thr Ala Ser Ala Asn Tyr Ser Lys Ala Val His His Val Leu Asp
465                 470                 475                 480
Val Ile His Glu Val Leu His His Gln Arg His Val Arg Thr Ile Trp
                485                 490                 495
Gln His Arg Lys Val Arg Leu His Gln Arg Leu Gln Leu Cys Val Phe
            500                 505                 510
Gln Gln Glu Val Gln Gln Val Leu Asp Trp Ile Glu Asn His Gly Glu
        515                 520                 525
Ala Phe Leu Ser Lys His Thr Gly Val Gly Lys Ser Leu His Arg Ala
    530                 535                 540
Arg Ala Leu Gln Lys Arg His Glu Asp Phe Glu Glu Val Ala Gln Asn
545                 550                 555                 560
Thr Tyr Thr Asn Ala Asp Lys Leu Leu Glu Ala Ala Glu Gln Leu Ala
                565                 570                 575
Gln Thr Gly Glu Cys Asp Pro Glu Glu Ile Tyr Gln Ala Ala His Gln
            580                 585                 590
Leu Glu Asp Arg Ile Gln Asp Phe Val Arg Arg Val Glu Gln Arg Lys
        595                 600                 605
Ile Leu Leu Asp Met Ser Val Ser Phe His Thr His Val Lys Glu Leu
    610                 615                 620
Trp Thr Trp Leu Glu Glu Leu Gln Lys Glu Leu Leu Asp Asp Val Tyr
625                 630                 635                 640
Ala Glu Ser Val Glu Ala Val Gln Asp Leu Ile Lys Arg Phe Gly Gln
                645                 650                 655
Gln Gln Gln Thr Thr Leu Gln Val Thr Val Asn Val Ile Lys Glu Gly
            660                 665                 670
Glu Asp Leu Ile Gln Gln Leu Arg Asp Ser Ala Ile Ser Ser Asn Lys
        675                 680                 685
Thr Pro His Asn Ser Ser Ile Asn His Ile Glu Thr Val Leu Gln Gln
    690                 695                 700
```

Leu Asp Glu Ala Gln Ser Gln Met Glu Glu Leu Phe Gln Glu Arg Lys
705 710 715 720
Ile Lys Leu Glu Leu Phe Leu His Val Arg Ile Phe Glu Arg Asp Ala
725 730 735
Ile Asp Ile Ile Ser Asp Leu Glu Ser Trp Asn Asp Glu Leu Ser Gln
740 745 750
Gln Met Asn Asp Phe Asp Thr Glu Asp Leu Thr Ile Ala Glu Gln Arg
755 760 765
Leu Gln His His Ala Asp Lys Ala Leu Thr Met Asn Asn Leu Thr Phe
770 775 780
Asp Val Ile His Gln Gly Gln Asp Leu Leu Gln Tyr Val Asn Glu Val
785 790 795 800
Gln Ala Ser Gly Val Glu Leu Leu Cys Asp Arg Asp Val Asp Met Ala
805 810 815
Thr Arg Val Gln Asp Leu Leu Glu Phe Leu His Glu Lys Gln Gln Glu
820 825 830
Leu Asp Leu Ala Ala Glu Gln His Arg Lys His Leu Glu Gln Cys Val
835 840 845
Gln Leu Arg His Leu Gln Ala Glu Val Lys Gln Val Leu Gly Trp Ile
850 855 860
Arg Asn Gly Glu Ser Met Leu Asn Ala Gly Leu Ile Thr Ala Ser Ser
865 870 875 880
Leu Gln Glu Ala Glu Gln Leu Gln Arg Glu His Glu Gln Phe Gln His
885 890 895
Ala Ile Glu Lys Thr His Gln Ser Ala Leu Gln Val Gln Gln Lys Ala
900 905 910
Glu Ala Met Leu Gln Ala Asn His Tyr Asp Met Asp Met Ile Arg Asp
915 920 925
Cys Ala Glu Lys Val Ala Ser His Trp Gln Gln Leu Met Leu Lys Met
930 935 940
Glu Asp Arg Leu Lys Leu Val Asn Ala Ser Val Ala Phe Tyr Lys Thr
945 950 955 960
Ser Glu Gln Val Cys Ser Val Leu Glu Ser Leu Glu Gln Glu Tyr Lys
965 970 975
Arg Glu Glu Asp Trp Cys Gly Gly Ala Asp Lys Leu Gly Pro Asn Ser
980 985 990
Glu Thr Asp His Val Thr Pro Met Ile Ser Lys His Leu Glu Gln Lys
995 1000 1005
Glu Ala Phe Leu Lys Ala Cys Thr Leu Ala Arg Arg Asn Ala Asp Val
1010 1015 1020
Phe Leu Lys Tyr Leu His Arg Asn Ser Val Asn Met Pro Gly Met Val
1025 1030 1035 1040
Thr His Ile Lys Ala Pro Glu Gln Gln Val Lys Asn Ile Leu Asn Glu
1045 1050 1055
Leu Phe Gln Arg Glu Asn Arg Val Leu His Tyr Trp Thr Met Arg Lys
1060 1065 1070
Arg Arg Leu Asp Gln Cys Gln Gln Tyr Val Val Phe Glu Arg Ser Ala
1075 1080 1085
Lys Gln Ala Leu Glu Trp Ile His Asp Asn Gly Glu Phe Tyr Leu Ser
1090 1095 1100
Thr His Thr Ser Thr Gly Ser Ser Ile Gln His Thr Gln Glu Leu Leu
1105 1110 1115 1120
Lys Glu His Glu Glu Phe Gln Ile Thr Ala Lys Gln Thr Lys Glu Arg
1125 1130 1135
Val Lys Leu Leu Ile Gln Leu Ala Asp Gly Phe Cys Glu Lys Gly His
1140 1145 1150
Ala His Ala Ala Glu Ile Lys Lys Cys Val Thr Ala Val Asp Lys Arg

1155 1160 1165
Tyr Arg Asp Phe Ser Leu Arg Met Glu Lys Tyr Arg Thr Ser Leu Glu
1170 1175 1180
Lys Ala Leu Gly Ile Ser Ser Asp Ser Asn Lys Ser Ser Lys Ser Leu
1185 1190 1195 1200
Gln Leu Asp Ile Ile Pro Ala Ser Ile Pro Gly Ser Glu Val Lys Leu
1205 1210 1215
Arg Asp Ala Ala His Glu Leu Asn Glu Glu Lys Arg Lys Ser Ala Arg
1220 1225 1230
Arg Lys Glu Phe Ile Met Ala Glu Leu Ile Gln Thr Glu Lys Ala Tyr
1235 1240 1245
Val Arg Asp Leu Arg Glu Cys Met Asp Thr Tyr Leu Trp Glu Met Thr
1250 1255 1260
Ser Gly Val Glu Glu Ile Pro Pro Gly Ile Val Asn Lys Glu Leu Ile
1265 1270 1275 1280
Ile Phe Gly Asn Met Gln Glu Ile Tyr Glu Phe His Asn Asn Ile Phe
1285 1290 1295
Leu Lys Glu Leu Glu Lys Tyr Glu Gln Leu Pro Glu Asp Val Gly His
1300 1305 1310
Cys Phe Val Thr Trp Ala Asp Lys Phe Gln Met Tyr Val Thr Tyr Cys
1315 1320 1325
Lys Asn Lys Pro Asp Ser Thr Gln Leu Ile Leu Glu His Ala Gly Ser
1330 1335 1340
Tyr Phe Asp Glu Ile Gln Gln Arg His Gly Leu Ala Asn Ser Ile Ser
1345 1350 1355 1360
Ser Tyr Leu Ile Lys Pro Val Gln Arg Ile Thr Lys Tyr Gln Leu Leu
1365 1370 1375
Leu Lys Glu Leu Leu Thr Cys Cys Glu Glu Gly Lys Gly Glu Ile Lys
1380 1385 1390
Asp Gly Leu Glu Val Met Leu Ser Val Pro Lys Arg Ala Asn Asp Ala
1395 1400 1405
Met His Leu Ser Met Leu Glu Gly Phe Asp Glu Asn Ile Glu Ser Gln
1410 1415 1420
Gly Glu Leu Ile Leu Gln Glu Ser Phe Gln Val Trp Asp Pro Lys Thr
1425 1430 1435 1440
Leu Ile Arg Lys Gly Arg Glu Arg His Leu Phe Leu Phe Glu Met Ser
1445 1450 1455
Leu Val Phe Ser Lys Glu Val Lys Asp Ser Ser Gly Arg Ser Lys Tyr
1460 1465 1470
Leu Tyr Lys Ser Lys Leu Phe Thr Ser Glu Leu Gly Val Thr Glu His
1475 1480 1485
Val Glu Gly Asp Pro Cys Lys Phe Ala Leu Trp Val Gly Arg Thr Pro
1490 1495 1500
Thr Ser Asp Asn Lys Ile Val Leu Lys Ala Ser Ser Ile Glu Asn Lys
1505 1510 1515 1520
Gln Asp Trp Ile Lys His Ile Arg Glu Val Ile Gln Glu Arg Thr Ile
1525 1530 1535
His Leu Lys Gly Ala Leu Lys Glu Pro Ile His Ile Pro Lys Thr Ala
1540 1545 1550
Pro Ala Thr Arg Gln Lys Gly Arg Arg Asp Gly Glu Asp Leu Asp Ser
1555 1560 1565
Gln Gly Asp Gly Ser Ser Gln Pro Asp Thr Ile Ser Ile Ala Ser Arg
1570 1575 1580
Thr Ser Gln Asn Thr Leu Asp Ser Asp Lys Leu Ser Gly Gly Cys Glu
1585 1590 1595 1600
Leu Thr Val Val Ile His Asp Phe Thr Ala Cys Asn Ser Asn Glu Leu
1605 1610 1615

Thr Ile Arg Arg Gly Gln Thr Val Glu Val Leu Glu Arg Pro His Asp
1620 1625 1630
Lys Pro Asp Trp Cys Leu Val Arg Thr Thr Asp Arg Ser Pro Ala Ala
1635 1640 1645
Glu Gly Leu Val Pro Cys Gly Ser Leu Cys Ile Ala His Ser Arg Ser
1650 1655 1660
Ser Met Glu Met Glu Gly Ile Phe Asn His Lys Asp Ser Leu Ser Val
1665 1670 1675 1680
Ser Ser Asn Asp Ala Ser Pro Pro Ala Ser Val Ala Ser Leu Gln Pro
1685 1690 1695
His Met Ile Gly Ala Gln Ser Ser Pro Gly Pro Lys Arg Pro Gly Asn
1700 1705 1710
Thr Leu Arg Lys Trp Leu Thr Ser Pro Val Arg Arg Leu Ser Ser Gly
1715 1720 1725
Lys Ala Asp Gly His Val Lys Lys Leu Ala His Lys His Lys Lys Ser
1730 1735 1740
Arg Glu Val Arg Lys Ser Ala Asp Ala Gly Ser Gln Lys Asp Ser Asp
1745 1750 1755 1760
Asp Ser Ala Ala Thr Pro Gln Asp Glu Thr Val Glu Glu Arg Gly Arg
1765 1770 1775
Asn Glu Gly Leu Ser Ser Gly Thr Leu Ser Lys Ser Ser Ser Ser Gly
1780 1785 1790
Met Gln Ser Cys Gly Glu Glu Glu Gly Glu Glu Gly Ala Asp Ala Val
1795 1800 1805
Pro Leu Pro Pro Pro Met Ala Ile Gln Gln His Ser Leu Leu Gln Pro
1810 1815 1820
Asp Ser Gln Asp Asp Lys Ala Ser Ser Arg Leu Leu Val Arg Pro Thr
1825 1830 1835 1840
Ser Ser Glu Thr Pro Ser Ala Ala Glu Leu Val Ser Ala Ile Glu Glu
1845 1850 1855
Leu Val Lys Ser Lys Met Ala Leu Glu Asp Arg Pro Ser Ser Leu Leu
1860 1865 1870
Val Asp Gln Gly Asp Ser Ser Ser Pro Ser Phe Asn Pro Ser Asp Asn
1875 1880 1885
Ser Leu Leu Ser Ser Ser Ser Pro Ile Asp Glu Met Glu Glu Arg Lys
1890 1895 1900
Ser Ser Ser Leu Lys Arg Arg His Tyr Val Leu Gln Glu Leu Val Glu
1905 1910 1915 1920
Thr Glu Arg Asp Tyr Val Arg Asp Leu Gly Tyr Val Val Glu Gly Tyr
1925 1930 1935
Met Ala Leu Met Lys Glu Asp Gly Val Pro Asp Asp Met Lys Gly Lys
1940 1945 1950
Asp Lys Ile Val Phe Gly Asn Ile His Gln Ile Tyr Asp Trp His Arg
1955 1960 1965
Asp Phe Phe Leu Gly Glu Leu Glu Lys Cys Leu Glu Asp Pro Glu Lys
1970 1975 1980
Leu Gly Ser Leu Phe Val Lys His Glu Arg Arg Leu His Met Tyr Ile
1985 1990 1995 2000
Ala Tyr Cys Gln Asn Lys Pro Lys Ser Glu His Ile Val Ser Glu Tyr
2005 2010 2015
Ile Asp Thr Phe Phe Glu Asp Leu Lys Gln Arg Leu Gly His Arg Leu
2020 2025 2030
Gln Leu Thr Asp Leu Leu Ile Lys Pro Val Gln Arg Ile Met Lys Tyr
2035 2040 2045
Gln Leu Leu Leu Lys Asp Phe Leu Lys Tyr Ser Lys Lys Ala Ser Leu
2050 2055 2060
Asp Thr Ser Glu Leu Glu Arg Ala Val Glu Val Met Cys Ile Val Pro

2065 2070 2075 2080
Arg Arg Cys Asn Asp Met Met Asn Val Gly Arg Leu Gln Gly Phe Asp
2085 2090 2095
Gly Lys Ile Val Ala Gln Gly Lys Leu Leu Leu Gln Asp Thr Phe Leu
2100 2105 2110
Val Thr Asp Gln Asp Ala Gly Leu Leu Pro Arg Cys Arg Glu Arg Arg
2115 2120 2125
Ile Phe Leu Phe Glu Gln Ile Val Ile Phe Ser Glu Pro Leu Asp Lys
2130 2135 2140
Lys Lys Gly Phe Ser Met Pro Gly Phe Leu Phe Lys Asn Ser Ile Lys
2145 2150 2155 2160
Val Ser Cys Leu Cys Leu Glu Glu Asn Val Glu Asn Asp Pro Cys Lys
2165 2170 2175
Phe Ala Leu Thr Ser Arg Thr Gly Asp Val Val Glu Thr Phe Ile Leu
2180 2185 2190
His Ser Ser Ser Pro Ser Val Arg Gln Thr Trp Ile His Glu Ile Asn
2195 2200 2205
Gln Ile Leu Glu Asn Gln Arg Asn Phe Leu Asn Ala Leu Thr Ser Pro
2210 2215 2220
Ile Glu Tyr Gln Arg Asn His Ser Gly Gly Gly Gly Gly Gly Gly Ser
2225 2230 2235 2240
Gly Ala Ala Ala Gly Val Gly Ala Ala Ala Ala Ala Gly Pro Pro Val
2245 2250 2255
Ala Ala Ala Ala Thr Val Ala Ala Pro Ala Ala Ala Ala Ala Pro Pro
2260 2265 2270
Ala Arg Ala Gly Ala Gly Pro Pro Gly Ser Pro Ser Leu Ser Asp Thr
2275 2280 2285
Thr Pro Pro Cys Trp Ser Pro Leu Gln Pro Arg Ala Arg Gln Arg Gln
2290 2295 2300
Thr Arg Cys Gln Ser Glu Ser Ser Ser Ser Ser Asn Ile Ser Thr Met
2305 2310 2315 2320
Leu Val Thr His Asp Tyr Thr Ala Val Lys Glu Asp Glu Ile Asn Val
2325 2330 2335
Tyr Gln Gly Glu Val Val Gln Ile Leu Ala Ser Asn Gln Gln Asn Met
2340 2345 2350
Phe Leu Val Phe Arg Ala Ala Thr Asp Gln Cys Pro Ala Ala Glu Gly
2355 2360 2365
Trp Ile Pro Gly Phe Val Leu Gly His Thr Ser Ala Val Ile Val Glu
2370 2375 2380
Asn Pro Asp Gly Thr Leu Lys Lys Ser Thr Ser Trp His Thr Ala Leu
2385 2390 2395 2400
Arg Leu Arg Lys Lys Ser Glu Lys Lys Asp Lys Asp Gly Lys Arg Glu
2405 2410 2415
Gly Lys Leu Glu Asn Gly Tyr Arg Lys Ser Arg Glu Gly Leu Ser Asn
2420 2425 2430
Lys Val Ser Val Lys Leu Leu Asn Pro Asn Tyr Ile Tyr Asp Val Pro
2435 2440 2445
Pro Glu Phe Val Ile Pro Leu Ser Glu Val Thr Cys Glu Thr Gly Glu
2450 2455 2460
Thr Val Val Leu Arg Cys Arg Val Cys Gly Arg Pro Lys Ala Ser Ile
2465 2470 2475 2480
Thr Trp Lys Gly Pro Glu His Asn Thr Leu Asn Asn Asp Gly His Tyr
2485 2490 2495
Ser Ile Ser Tyr Ser Asp Leu Gly Glu Ala Thr Leu Lys Ile Val Gly
2500 2505 2510
Val Thr Thr Glu Asp Asp Gly Ile Tyr Thr Cys Ile Ala Val Asn Asp
2515 2520 2525

Met Gly Ser Ala Ser Ser Ser Ala Ser Leu Arg Val Leu Gly Pro Gly
2530 2535 2540
Met Asp Gly Ile Met Val Thr Trp Lys Asp Asn Phe Asp Ser Phe Tyr
2545 2550 2555 2560
Ser Glu Val Ala Glu Leu Gly Arg Gly Arg Phe Ser Val Val Lys Lys
2565 2570 2575
Cys Asp Gln Lys Gly Thr Lys Arg Ala Val Ala Thr Lys Phe Val Asn
2580 2585 2590
Lys Lys Leu Met Lys Arg Asp Gln Val Thr His Glu Leu Gly Ile Leu
2595 2600 2605
Gln Ser Leu Gln His Pro Leu Leu Val Gly Leu Leu Asp Thr Phe Glu
2610 2615 2620
Thr Pro Thr Ser Tyr Ile Leu Val Leu Glu Met Ala Asp Gln Gly Arg
2625 2630 2635 2640
Leu Leu Asp Cys Val Val Arg Trp Gly Ser Leu Thr Glu Gly Lys Ile
2645 2650 2655
Arg Ala His Leu Gly Glu Val Leu Glu Ala Val Arg Tyr Leu His Asn
2660 2665 2670
Cys Arg Ile Ala His Leu Asp Leu Lys Pro Glu Asn Ile Leu Val Asp
2675 2680 2685
Glu Ser Leu Ala Lys Pro Thr Ile Lys Leu Ala Asp Phe Gly Asp Ala
2690 2695 2700
Val Gln Leu Asn Thr Thr Tyr Tyr Ile His Gln Leu Leu Gly Asn Pro
2705 2710 2715 2720
Glu Phe Ala Ala Pro Glu Ile Ile Leu Gly Asn Pro Val Ser Leu Thr
2725 2730 2735
Ser Asp Thr Trp Ser Val Gly Val Leu Thr Tyr Val Leu Leu Ser Gly
2740 2745 2750
Val Ser Pro Phe Leu Asp Asp Ser Val Glu Glu Thr Cys Leu Asn Ile
2755 2760 2765
Cys Arg Leu Asp Phe Ser Phe Pro Asp Asp Tyr Phe Lys Gly Val Ser
2770 2775 2780
Gln Lys Ala Lys Glu Phe Val Cys Phe Leu Leu Gln Glu Asp Pro Ala
2785 2790 2795 2800
Lys Arg Pro Ser Ala Ala Leu Ala Leu Gln Glu Gln Trp Leu Gln Ala
2805 2810 2815
Gly Asn Gly Arg Ser Thr Gly Val Leu Asp Thr Ser Arg Leu Thr Ser
2820 2825 2830
Phe Ile Glu Arg Arg Lys His Gln Asn Asp Val Arg Pro Ile Arg Ser
2835 2840 2845
Ile Lys Asn Phe Leu Gln Ser Arg Leu Leu Pro Arg Val
2850 2855 2860

<210> 109
<211> 271
<212> PRT
<213> Homo sapiens

<400> 109
Met Val Leu Ile Lys Glu Phe Arg Val Val Leu Pro Cys Ser Val Gln
1 5 10 15
Glu Tyr Gln Val Gly Gln Leu Tyr Ser Val Ala Glu Ala Ser Lys Asn
20 25 30
Glu Thr Gly Gly Gly Glu Gly Ile Glu Val Leu Lys Asn Glu Pro Tyr
35 40 45
Glu Lys Asp Gly Glu Lys Gly Gln Tyr Thr His Lys Ile Tyr His Leu
50 55 60

Lys Ser Lys Val Pro Ala Phe Val Arg Met Ile Ala Pro Glu Gly Ser
65 70 75 80
Leu Val Phe His Glu Lys Ala Trp Asn Ala Tyr Pro Tyr Cys Arg Thr
85 90 95
Ile Val Thr Asn Glu Tyr Met Lys Asp Asp Phe Phe Ile Lys Ile Glu
100 105 110
Thr Trp His Lys Pro Asp Leu Gly Thr Leu Glu Asn Val His Gly Leu
115 120 125
Asp Pro Asn Thr Trp Lys Thr Val Glu Ile Val His Ile Asp Ile Ala
130 135 140
Asp Arg Ser Gln Val Glu Pro Ala Asp Tyr Lys Ala Asp Glu Asp Pro
145 150 155 160
Ala Leu Phe Gln Ser Val Lys Thr Lys Arg Gly Pro Leu Gly Pro Asn
165 170 175
Trp Lys Lys Glu Leu Ala Asn Ser Pro Asp Cys Pro Gln Met Cys Ala
180 185 190
Tyr Lys Leu Val Thr Ile Lys Phe Lys Trp Trp Gly Leu Gln Ser Lys
195 200 205
Val Glu Asn Phe Ile Gln Lys Gln Glu Lys Arg Ile Phe Thr Asn Phe
210 215 220
His Arg Gln Leu Phe Cys Trp Ile Asp Lys Trp Ile Asp Leu Thr Met
225 230 235 240
Glu Asp Ile Arg Arg Met Glu Asp Glu Thr Gln Lys Glu Leu Glu Thr
245 250 255
Met Arg Lys Arg Gly Ser Val Arg Gly Thr Ser Ala Ala Asp Val
260 265 270

<210> 110
<211> 233
<212> PRT
<213> Homo sapiens

<400> 110
Asn Ser Val Leu Asn Ser Asn Ala Ile Lys Asn Leu Pro Pro Pro Leu
1 5 10 15
Gly Gly Ala Ala Gly His Pro Gly Ser Ala Val Ser Ala Ala Pro Gly
20 25 30
Ile Leu Tyr Pro Gly Gly Asn Lys Tyr Gln Thr Ile Asp Asn Tyr Gln
35 40 45
Pro Tyr Pro Cys Ala Glu Asp Glu Glu Cys Gly Thr Asp Glu Tyr Cys
50 55 60
Ala Ser Pro Thr Arg Gly Gly Asp Ala Gly Val Gln Ile Cys Leu Ala
65 70 75 80
Cys Arg Lys Arg Arg Lys Arg Cys Met Arg His Ala Met Cys Cys Pro
85 90 95
Gly Asn Tyr Cys Lys Asn Gly Ile Cys Val Ser Ser Asp Gln Asn His
100 105 110
Phe Arg Gly Glu Ile Glu Glu Thr Ile Thr Glu Ser Phe Gly Asn Asp
115 120 125
His Ser Thr Leu Asp Gly Tyr Ser Arg Arg Thr Thr Leu Ser Ser Lys
130 135 140
Met Tyr His Thr Lys Gly Gln Glu Gly Ser Val Cys Leu Arg Ser Ser
145 150 155 160
Asp Cys Ala Ser Gly Leu Cys Cys Ala Arg His Phe Trp Ser Lys Ile
165 170 175
Cys Lys Pro Val Leu Lys Glu Gly Gln Val Cys Thr Lys His Arg Arg
180 185 190

Lys Gly Ser His Gly Leu Glu Ile Phe Gln Arg Cys Tyr Cys Gly Glu
195 200 205
Gly Leu Ser Cys Arg Ile Gln Lys Asp His His Gln Ala Ser Asn Ser
210 215 220
Ser Arg Leu His Thr Cys Gln Arg His
225 230

<210> 111
<211> 212
<212> PRT
<213> Homo sapiens

<400> 111
Asn Ser Val Leu Asn Ser Asn Ala Ile Lys Asn Leu Pro Pro Pro Leu
1 5 10 15
Gly Gly Ala Ala Gly His Pro Gly Ser Ala Val Ser Ala Ala Pro Gly
20 25 30
Ile Leu Tyr Pro Gly Gly Asn Lys Tyr Gln Thr Ile Asp Asn Tyr Gln
35 40 45
Pro Tyr Pro Cys Ala Glu Asp Glu Glu Cys Gly Thr Asp Glu Tyr Cys
50 55 60
Ala Ser Pro Thr Arg Gly Gly Asp Ala Gly Val Gln Ile Cys Leu Ala
65 70 75 80
Cys Arg Lys Arg Arg Lys Arg Cys Met Arg His Ala Met Cys Cys Pro
85 90 95
Gly Asn Tyr Cys Lys Asn Gly Ile Cys Val Ser Ser Asp Gln Asn His
100 105 110
Phe Arg Gly Glu Ile Glu Glu Thr Ile Thr Glu Ser Phe Gly Asn Asp
115 120 125
His Ser Thr Leu Asp Gly Tyr Ser Arg Arg Thr Thr Leu Ser Ser Lys
130 135 140
Met Tyr His Thr Lys Gly Gln Glu Gly Ser Val Cys Leu Arg Ser Ser
145 150 155 160
Asp Cys Ala Ser Gly Leu Cys Cys Ala Arg His Phe Trp Ser Lys Ile
165 170 175
Cys Lys Pro Val Leu Lys Glu Gly Gln Val Cys Thr Lys His Arg Arg
180 185 190
Lys Gly Ser His Gly Leu Glu Ile Phe Gln Arg Cys Tyr Cys Gly Glu
195 200 205
Gly Leu Ser Cys
210

<210> 112
<211> 149
<212> PRT
<213> Homo sapiens

<400> 112
Asn Ser Val Leu Asn Ser Asn Ala Ile Lys Asn Leu Pro Pro Pro Leu
1 5 10 15
Gly Gly Ala Ala Gly His Pro Gly Ser Ala Val Ser Ala Ala Pro Gly
20 25 30
Ile Leu Tyr Pro Gly Gly Asn Lys Tyr Gln Thr Ile Asp Asn Tyr Gln
35 40 45
Pro Tyr Pro Cys Ala Glu Asp Glu Glu Cys Gly Thr Asp Glu Tyr Cys

```
    50                  55                  60
Ala Ser Pro Thr Arg Gly Gly Asp Ala Gly Val Gln Ile Cys Leu Ala
65                  70                  75                  80
Cys Arg Lys Arg Arg Lys Arg Cys Met Arg His Ala Met Cys Cys Pro
                85                  90                  95
Gly Asn Tyr Cys Lys Asn Gly Ile Cys Val Ser Ser Asp Gln Asn His
            100                 105                 110
Phe Arg Gly Glu Ile Glu Glu Thr Ile Thr Glu Ser Phe Gly Asn Asp
        115                 120                 125
His Ser Thr Leu Asp Gly Tyr Ser Arg Arg Thr Thr Leu Ser Ser Lys
    130                 135                 140
Met Tyr His Thr Lys
145
```

<210> 113
<211> 170
<212> PRT
<213> Homo sapiens

<400> 113

```
Cys Ala Ser Pro Thr Arg Gly Gly Asp Ala Gly Val Gln Ile Cys Leu
1               5                   10                  15
Ala Cys Arg Lys Arg Arg Lys Arg Cys Met Arg His Ala Met Cys Cys
            20                  25                  30
Pro Gly Asn Tyr Cys Lys Asn Gly Ile Cys Val Ser Ser Asp Gln Asn
        35                  40                  45
His Phe Arg Gly Glu Ile Glu Glu Thr Ile Thr Glu Ser Phe Gly Asn
    50                  55                  60
Asp His Ser Thr Leu Asp Gly Tyr Ser Arg Arg Thr Thr Leu Ser Ser
65                  70                  75                  80
Lys Met Tyr His Thr Lys Gly Gln Glu Gly Ser Val Cys Leu Arg Ser
                85                  90                  95
Ser Asp Cys Ala Ser Gly Leu Cys Cys Ala Arg His Phe Trp Ser Lys
            100                 105                 110
Ile Cys Lys Pro Val Leu Lys Glu Gly Gln Val Cys Thr Lys His Arg
        115                 120                 125
Arg Lys Gly Ser His Gly Leu Glu Ile Phe Gln Arg Cys Tyr Cys Gly
    130                 135                 140
Glu Gly Leu Ser Cys Arg Ile Gln Lys Asp His His Gln Ala Ser Asn
145                 150                 155                 160
Ser Ser Arg Leu His Thr Cys Gln Arg His
                165                 170
```

<210> 114
<211> 128
<212> PRT
<213> Homo sapiens

<400> 114

```
Val Ser Ser Asp Gln Asn His Phe Arg Gly Glu Ile Glu Glu Thr Ile
1               5                   10                  15
Thr Glu Ser Phe Gly Asn Asp His Ser Thr Leu Asp Gly Tyr Ser Arg
            20                  25                  30
Arg Thr Thr Leu Ser Ser Lys Met Tyr His Thr Lys Gly Gln Glu Gly
        35                  40                  45
```

Ser Val Cys Leu Arg Ser Ser Asp Cys Ala Ser Gly Leu Cys Cys Ala
50 55 60
Arg His Phe Trp Ser Lys Ile Cys Lys Pro Val Leu Lys Glu Gly Gln
65 70 75 80
Val Cys Thr Lys His Arg Arg Lys Gly Ser His Gly Leu Glu Ile Phe
85 90 95
Gln Arg Cys Tyr Cys Gly Glu Gly Leu Ser Cys Arg Ile Gln Lys Asp
100 105 110
His His Gln Ala Ser Asn Ser Ser Arg Leu His Thr Cys Gln Arg His
115 120 125

<210> 115
<211> 84
<212> PRT
<213> Homo sapiens

<400> 115
Gly Gln Glu Gly Ser Val Cys Leu Arg Ser Ser Asp Cys Ala Ser Gly
1 5 10 15
Leu Cys Cys Ala Arg His Phe Trp Ser Lys Ile Cys Lys Pro Val Leu
20 25 30
Lys Glu Gly Gln Val Cys Thr Lys His Arg Arg Lys Gly Ser His Gly
35 40 45
Leu Glu Ile Phe Gln Arg Cys Tyr Cys Gly Glu Gly Leu Ser Cys Arg
50 55 60
Ile Gln Lys Asp His His Gln Ala Ser Asn Ser Ser Arg Leu His Thr
65 70 75 80
Cys Gln Arg His

<210> 116
<211> 149
<212> PRT
<213> Homo sapiens
<400> 116
Cys Ala Ser Pro Thr Arg Gly Gly Asp Ala Gly Val Gln Ile Cys Leu
1 5 10 15
Ala Cys Arg Lys Arg Arg Lys Arg Cys Met Arg His Ala Met Cys Cys
20 25 30
Pro Gly Asn Tyr Cys Lys Asn Gly Ile Cys Val Ser Ser Asp Gln Asn
35 40 45
His Phe Arg Gly Glu Ile Glu Glu Thr Ile Thr Glu Ser Phe Gly Asn
50 55 60
Asp His Ser Thr Leu Asp Gly Tyr Ser Arg Arg Thr Thr Leu Ser Ser
65 70 75 80
Lys Met Tyr His Thr Lys Gly Gln Glu Gly Ser Val Cys Leu Arg Ser
85 90 95
Ser Asp Cys Ala Ser Gly Leu Cys Cys Ala Arg His Phe Trp Ser Lys
100 105 110
Ile Cys Lys Pro Val Leu Lys Glu Gly Gln Val Cys Thr Lys His Arg
115 120 125
Arg Lys Gly Ser His Gly Leu Glu Ile Phe Gln Arg Cys Tyr Cys Gly
130 135 140
Glu Gly Leu Ser Cys
145

<210> 117
<211> 107
<212> PRT
<213> Homo sapiens

<400> 117
```
Val Ser Ser Asp Gln Asn His Phe Arg Gly Glu Ile Glu Glu Thr Ile
1               5                   10                  15
Thr Glu Ser Phe Gly Asn Asp His Ser Thr Leu Asp Gly Tyr Ser Arg
            20                  25                  30
Arg Thr Thr Leu Ser Ser Lys Met Tyr His Thr Lys Gly Gln Glu Gly
        35                  40                  45
Ser Val Cys Leu Arg Ser Ser Asp Cys Ala Ser Gly Leu Cys Cys Ala
    50                  55                  60
Arg His Phe Trp Ser Lys Ile Cys Lys Pro Val Leu Lys Glu Gly Gln
65                  70                  75                  80
Val Cys Thr Lys His Arg Arg Lys Gly Ser His Gly Leu Glu Ile Phe
                85                  90                  95
Gln Arg Cys Tyr Cys Gly Glu Gly Leu Ser Cys
            100                 105
```

<210> 118
<211> 15
<212> PRT
<213> Homo sapiens

<400> 118
```
Gly Asn Lys Tyr Gln Thr Ile Asp Asn Tyr Gln Pro Tyr Pro Cys
1               5                   10                  15
```

<210> 119
<211> 22
<212> PRT
<213> Homo sapiens

<400> 119
```
Leu Asp Gly Tyr Ser Arg Arg Thr Thr Leu Ser Ser Lys Met Tyr His
1               5                   10                  15
Thr Lys Gly Gln Glu Gly
            20
```

<210> 120
<211> 21
<212> PRT
<213> Homo sapiens

<400> 120
```
Arg Ile Gln Lys Asp His His Gln Ala Ser Asn Ser Ser Arg Leu His
1               5                   10                  15
Thr Cys Gln Arg His
            20
```

<210> 121
<211> 15
<212> PRT
<213> Homo sapiens

<400> 121
Arg Gly Glu Ile Glu Glu Thr Ile Thr Glu Ser Phe Gly Asn Asp
1 5 10 15

<210> 122
<211> 19
<212> PRT
<213> Homo sapiens

<400> 122
Glu Ile Phe Gln Arg Cys Tyr Cys Gly Glu Gly Leu Ser Cys Arg Ile
1 5 10 15
Gln Lys Asp

<210> 123
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> LRP5/HBM amino acid sequence

<400> 123
Met Tyr Trp Thr Asp Trp Val Glu Thr Pro Arg Ile Glu
1 5 10

<210> 124
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> LRP5/HBM amino acid sequence

<400> 124
Met Tyr Trp Thr Asp Trp Gly Glu Thr Pro Arg Ile Glu
1 5 10

<210> 125
<211> 13
<212> PRT
<213> Artificial Sequence

<220>
<223> LRP5/HBM amino acid sequence

<400> 125
Lys Arg Thr Gly Gly Lys Arg Lys Glu Ile Leu Ser Ala
1 5 10

<210> 126
<211> 17
<212> PRT
<213> Artificial Sequence

<220>
<223> LRP5/HBM amino acid sequence

<400> 126
Glu Arg Val Glu Lys Thr Thr Gly Asp Lys Arg Thr Arg Ile Gln Gly
1 5 10 15
Arg

<210> 127
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> LRP5/HBM amino acid sequence

<400> 127
Lys Gln Gln Cys Asp Ser Phe Pro Asp Cys Ile Asp Gly Ser Asp Glu
1 5 10 15

<210> 128
<211> 266
<212> PRT
<213> Homo sapiens

<400> 128
Met Met Ala Leu Gly Ala Ala Gly Ala Thr Arg Val Phe Val Ala Met
1 5 10 15
Val Ala Ala Ala Leu Gly Gly His Pro Leu Leu Gly Val Ser Ala Thr
20 25 30
Leu Asn Ser Val Leu Asn Ser Asn Ala Ile Lys Asn Leu Pro Pro Pro
35 40 45
Leu Gly Gly Ala Ala Gly His Pro Gly Ser Ala Val Ser Ala Ala Pro
50 55 60
Gly Ile Leu Tyr Pro Gly Gly Asn Lys Tyr Gln Thr Ile Asp Asn Tyr
65 70 75 80
Gln Pro Tyr Pro Cys Ala Glu Asp Glu Glu Cys Gly Thr Asp Glu Tyr
85 90 95
Cys Ala Ser Pro Thr Arg Gly Gly Asp Ala Gly Val Gln Ile Cys Leu
100 105 110
Ala Cys Arg Lys Arg Arg Lys Arg Cys Met Arg His Ala Met Cys Cys
115 120 125
Pro Gly Asn Tyr Cys Lys Asn Gly Ile Cys Val Ser Ser Asp Gln Asn
130 135 140

```
His Phe Arg Gly Glu Ile Glu Glu Thr Ile Thr Glu Ser Phe Gly Asn
145                 150                 155                 160
Asp His Ser Thr Leu Asp Gly Tyr Ser Arg Arg Thr Thr Leu Ser Ser
                165                 170                 175
Lys Met Tyr His Thr Lys Gly Gln Glu Gly Ser Val Cys Leu Arg Ser
            180                 185                 190
Ser Asp Cys Ala Ser Gly Leu Cys Cys Ala Arg His Phe Trp Ser Lys
        195                 200                 205
Ile Cys Lys Pro Val Leu Lys Glu Gly Gln Val Cys Thr Lys His Arg
    210                 215                 220
Arg Lys Gly Ser His Gly Leu Glu Ile Phe Gln Arg Cys Tyr Cys Gly
225                 230                 235                 240
Glu Gly Leu Ser Cys Arg Ile Gln Lys Asp His His Gln Ala Ser Asn
                245                 250                 255
Ser Ser Arg Leu His Thr Cys Gln Arg His
            260                 265
```

<210> 129
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 129
tggtcagcgg cctggaggat gtggccgcag tggacttcc 39

<210> 130
<211> 39
<212> DNA
<213> Artificial Sequence
<220>
<223> complementary synthetic oligonucleotide

<400> 130
ggaagtccac tgcggccaca tcctccaggc cgctgacca 39

<210> 131
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 131
aagctgtact ggacggactc agtgaccaac cgcatcgagg 40

<210> 132
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 132
cctcgatgcg gttggtcact gagtccgtcc agtacagctt 40

<210> 133
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 133
atgtactgga cagactggaa ggagacgccc cggattgagc g 41

<210> 134
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 134
cgctcaatcc ggggcgtctc cttccagtct gtccagtaca t 41

<210> 135
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 135
atgtactgga cagactggtt tgagacgccc cggattgagc g 41

<210> 136
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 136
cgctcaatcc ggggcgtctc aaaccagtct gtccagtaca t 41

<210> 137
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 137

atgtactgga cagactggat tgagacgccc cggattgagc g 41

<210> 138
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide
<400> 138
cgctcaatcc ggggcgtctc aatccagtct gtccagtaca t 41

<210> 139
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 139
atgtactgga cagactggca ggagacgccc cggattgagc g 41

<210> 140
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 140
cgctcaatcc ggggcgtctc ctgccagtct gtccagtaca t 41

<210> 141
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 141
cggacattta ctggcccaat gtactgacca tcgacctgga gg 42

<210> 142
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 142
cctccaggtc gatggtcagt acattgggcc agtaaatgtc cg 42

<210> 143

<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 143
agctctactg ggctgacgtc aagctcagct tcatccaccg 40

<210> 144
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 144
cggtggatga agctgagctt gacgtcagcc cagtagagct 40

<210> 145
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 145
gagtgccctc tactcacccg tggacatcca ggtgctgagc c 41

<210> 146
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 146
ggctcagcac ctggatgtcc acgggtgagt agagggcact c 41
<210> 147
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 147
catgtactgg acagactggg tagagaaccc taaaatcgag tgtgc 45

<210> 148
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 148
gcacactcga ttttagggtt ctctacccag tctgtccagt acatg 45

<210> 149
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 149
catctactgg accgagtggg tcggcaagcc gaggatcgtg cg 42

<210> 150
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 150
cgcacgatcc tcggcttgcc gacccactcg gtccagtaga tg 42

<210> 151
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 151
gtacttcacc aacatggtgg accgggcagc caagatcgaa cg 42

<210> 152
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 152
cgttcgatct tggctgcccg gtccaccatg ttggtgaagt ac 42

<210> 153
<211> 42
<212> DNA
<213> Artificial Sequence

<220>

<223> complementary synthetic oligonucleotide

<400> 153
gtactggaca gactgggtag aagtgccaaa gatagaacgt gc 42

<210> 154
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> complementary synthetic oligonucleotide

<400> 154
gcacgttcta tctttggcac ttctacccag tctgtccagt ac 42

<210> 155
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 155
tttttgtcg accaattcca acgctatcaa g 31

<210> 156
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 156
tttttgtcg acctgcgcta gtcccacccg c 31

<210> 157
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 157
tttttgtcg accgtgtctt ctgatcaaaa tc 32

<210> 158
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 158
tttttgtcg accggacaag aaggttctgt ttg 33

<210> 159
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 159
ttttttgcgg ccgcttattt ggtgtgatac atttttg 37

<210> 160
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 160
ttttttgcgg ccgcttagca agacagacct tctcc 35

<210> 161
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 161
ttttttgcgg ccgcttagtg tctctgacaa gtgtg 35

<210> 162
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 162
cagtgaattc accatgcaaa acaccacttt gttc 34

<210> 163
<211> 32
<212> DNA
<213> Artificial Sequence
<220>
<223> primer

<400> 163
cagttgcggc cgctcatctc cggtggcctc tg 32

<210> 164
<211> 39
<212> DNA
<213> Homo sapiens

<220>
<223> primer

<400> 164
caatagtcga cgaattcacc atggctctgg gcgcagcgg 39

<210> 165
<211> 42
<212> DNA
<213> Homo sapiens

<220>
<223> primer

<400> 165
gtattgcggc cgctctagat tagtgtctct gacaagtgtg aa 42

<210> 166
<211> 32
<212> DNA
<213> Homo sapiens

<220>
<223> primer

<400> 166
ctaacggatc caccatggcc gcgttgatgc gg 32

<210> 167
<211> 33
<212> DNA
<213> Homo sapiens

<220>
<223> primer

<400> 167
gattcgaatt ctcaaatttt ctgacacaca tgg 33

<210> 168
<211> 690
<212> DNA
<213> Homo sapiens

<400> 168
aattccaacg ctatcaagaa cctgcccca ccgctgggcg gcgctgcggg gcacccaggc 60
tctgcagtca gcgccgcgcc gggaatcctg tacccggcg ggaataagta ccagaccatt 120
gacaactacc agccgtaccc gtgcgcagag gacgaggagt gcggcactga tgagtactgc 180
gctagtccca cccgcggagg ggacgcgggc gtgcaaatct gtctcgcctg caggaagcgc 240
cgaaaacgct gcatgcgtca cgctatgtgc tgccccggga attactgcaa aaatggaata 300
tgtgtgtctt ctgatcaaaa tcatttccga ggagaaattg aggaaaccat cactgaaagc 360
tttggtaatg atcatagcac cttggatggg tattccagaa gaaccacctt gtcttcaaaa 420

atgtatcaca ccaaaggaca agaaggttct gtttgtctcc ggtcatcaga ctgtgcctca 480
ggattgtgtt gtgctagaca cttctggtcc aagatctgta aacctgtcct gaaagaaggt 540
caagtgtgta ccaagcatag gagaaaaggc tctcatggac tagaaatatt ccagcgttgt 600
tactgtggag aaggtctgtc ttgccggata cagaaagatc accatcaagc cagtaattct 660
tctaggcttc acacttgtca gagacactaa 690

<210> 169
<211> 1226
<212> DNA
<213> Homo sapiens

<400> 169
ctcatcctgc ccctgcatgg actgaggaac gtcaaagcca tcgactatga cccactggac 60
aagttcatct actgggtgga tgggcgccag aacatcaagc gagccaagga cgacgggacc 120
cagccctttg ttttgacctc tctgagccaa ggccaaaacc cagacaggca gccccacgac 180
ctcagcatcg acatctacag ccggacactg ttctggacgt gcgaggccac caataccatc 240
aacgtccaca ggctgagcgg ggaagccatg gggtggtgc tgcgtgggga ccgcgacaag 300
cccagggcca tcgtcgtcaa cgcggagcga gggtacctgt acttcaccaa catgcaggac 360
cgggcagcca agatcgaacg cgcagccctg gacggcaccg agcgcgaggt cctcttcacc 420
accggcctca tccgccctgt ggccctggtg gtagacaaca cactgggcaa gctgttctgg 480
gtggacgcgg acctgaagcg cattgagagc tgtgacctgt caggggccaa ccgcctgacc 540
ctggaggacg ccaacatcgt gcagcctctg ggcctgacca tccttggcaa gcatctctac 600
tggatcgacc gccagcagca gatgatcgag cgtgtggaga agaccaccgg ggacaagcgg 660
actcgcatcc agggccgtgt cgcccacctc actggcatcc atgcagtgga ggaagtcagc 720
ctggaggagt tctcagccca cccatgtgcc cgtgacaatg gtggctgctc ccacatctgt 780
attgccaagg gtgatgggac accacggtgc tcatgcccag tccacctcgt gctcctgcag 840
aacctgctga cctgtggaga gccgcccacc tgctccccgg accagtttgc atgtgccaca 900
ggggagatcg actgtatccc cggggcctgg cgctgtgacg gctttcccga gtgcgatgac 960
cagagcgacg aggagggctg ccccgtgtgc tccgccgccc agttcccctg cgcgcggggt 1020
cagtgtggtg gacctgcgcc tgcgctgcga cggcgaggca gactgtcagg accgctcaga 1080
cgaggcggac tgtgacgcca tctgcctgcc caaccagttc cggtgtgcga gcggccagtg 1140
tgtcctcatc aaacagccag tgcgactcct tccccgactg tatcgacggc tccgacgagc 1200
tcatgtgtga aatcaccaag ccgccc 1226

<210> 170
<211> 934
<212> DNA
<213> Homo sapiens

<400> 170
agggccatcg tcgtcaacgc ggagcgaggg tacctgtact tcaccaacat gcaggaccgg 60
gcagccaaga tcgaacgcgc agccctggac ggcaccgagc gcgaggtcct cttcaccacc 120
ggcctcatcc gccctgtggc cctggtggta gacaacacac tgggcaagct gttctgggtg 180
gacgcggacc tgaagcgcat tgagagctgt gacctgtcag gggccaaccg cctgaccctg 240
gaggacgcca acatcgtgca gcctctgggc ctgaccatcc ttggcaagca tctctactgg 300
atcgaccgcc agcagcagat gatcgagcgt gtggagaaga ccaccgggga caagcggact 360
cgcatccagg gccgtgtcgc ccacctcact ggcatccatg cagtggagga agtcagcctg 420
gaggagttct cagcccaccc atgtgcccgt gacaatggtg gctgctccca catctgtatt 480
gccaagggtg atgggacacc acggtgctca tgcccagtcc acctcgtgct cctgcagaac 540
ctgctgacct gtggagagcc gcccacctgc tccccggacc agtttgcatg tgccacaggg 600
gagatcgact gtatccccgg ggcctggcgc tgtgacggct ttcccgagtg cgatgaccag 660
agcgacgagg agggctgccc cgtgtgcctc cgccgcccag ttcccctgcg cgcggggtca 720
gtgtgtggac ctgcgcctgc gctgcgacgg cgaggcagac tgtcaggacc gctcagacga 780
ggcggactgt gacgccatct ggcctgccca accagttccg gtgtgcgagc ggccagtgtg 840
tcctcatcaa acagcagtgc gactccttcc ccgactgtat cgacggctcc gacgagctca 900
tgtgtgaaat caccaagccg ccctaagcgg ccgc 934

<210> 171
<211> 16
<212> PRT
<213> Homo sapiens

<400> 171
Ser Val Gly Cys Leu Leu Cys Ala Gly Leu Gly Val Trp Ser Leu Ser
1 5 10 15

<210> 172
<211> 19
<212> PRT
<213> Homo sapiens

<400> 172
Trp Cys Cys Cys Gly Leu Phe Arg Gly Val Cys Val Trp Ser Cys Gly
1 5 10 15
Ala Asp Asp

<210> 173
<211> 16
<212> PRT
<213> Homo sapiens

<400> 173
Gly Trp Arg Arg Cys Asp Trp Cys Gly Cys Val Ser Trp Cys Trp Val
1 5 10 15

<210> 174
<211> 32
<212> PRT
<213> Homo sapiens

<400> 174
Met Pro Gly Ser Val Ser His Cys Trp Gly Gly Ile Cys Glu Ala Leu
1 5 10 15
Ser Cys Cys Ala Val Asp Val Cys Leu Arg Cys Gly Gly Trp Phe Arg
20 25 30

<210> 175
<211> 16
<212> PRT
<213> Homo sapiens

<400> 175
Ser Cys Cys Ala Val Asp Val Cys Leu Arg Cys Gly Gly Trp Phe Arg
1 5 10 15

<210> 176
<211> 16
<212> PRT

<213> Homo sapiens

<400> 176
Ser Val Leu Gly Thr Cys Cys Cys Cys Gly Gly Trp Ile Leu Cys Glu
1 5 10 15

<210> 177
<211> 16
<212> PRT
<213> Homo sapiens

<400> 177
Val Leu Ser Val Cys Glu Val Cys Gly Gly Val Phe Val Arg Arg Cys
1 5 10 15

<210> 178
<211> 15
<212> PRT
<213> Homo sapiens

<400> 178
Gly Met Trp Tyr Trp Ser Gly Arg Asp Cys Ala Leu Cys Trp Leu
1 5 10 15

<210> 179
<211> 16
<212> PRT
<213> Homo sapiens

<400> 179
Cys Thr Ala Val Met Trp Gly Val Gly Ser Val Ala Tyr Leu Gly Glu
1 5 10 15

<210> 180
<211> 13
<212> PRT
<213> Homo sapiens

<400> 180
Val Val Cys Trp Trp Cys Gly Cys Arg Gly Trp Trp Arg
1 5 10

<210> 181
<211> 16
<212> PRT
<213> Homo sapiens

<400> 181
Cys Val Cys Ala Ser Phe Cys Cys Cys Val Cys Gly Leu Arg Leu Leu
1 5 10 15

<210> 182
<211> 16
<212> PRT
<213> Homo sapiens

<400> 182

```
Thr Tyr Glu Val Cys Glu Glu Cys Gly Gly Arg Val Arg Met Trp Val
 1               5                  10                  15
```

<210> 183
<211> 16
<212> PRT
<213> Homo sapiens

<400> 183

```
Val Val Val Cys Ala Ser Cys Gly Gln Val Trp His Gly Ser Gly Ala
 1               5                  10                  15
```

<210> 184
<211> 16
<212> PRT
<213> Homo sapiens

<400> 184

```
Cys Cys Arg Cys Cys His Cys Trp Asp Cys Glu Trp His Met Cys Val
 1               5                  10                  15
```

<210> 185
<211> 16
<212> PRT
<213> Homo sapiens

<400> 185

```
Phe Cys Ala Ser Cys Cys Trp Cys Gly Cys Asp Cys Phe Gly Trp Val
 1               5                  10                  15
```

<210> 186
<211> 16
<212> PRT
<213> Homo sapiens

<400> 186

```
Cys Asp Tyr Cys Trp Ser Cys Gly Val Trp Cys Pro Ser Ser Trp Leu
 1               5                  10                  15
```

<210> 187
<211> 16
<212> PRT
<213> Homo sapiens

<400> 187

```
Val Tyr Leu Cys Val Trp Cys Gly Ala Ala Arg Phe Gly Cys Tyr Gly
```

1 5 10 15

<210> 188
<211> 16
<212> PRT
<213> Homo sapiens

<400> 188
Phe Cys Val Cys Gly Cys Cys Trp Cys Trp Cys Ala Ala Cys Trp Cys
1 5 10 15

<210> 189
<211> 16
<212> PRT
<213> Homo sapiens

<400> 189
Val Val Leu Cys Ser Arg Cys Gly Arg Leu Trp Arg Trp Ser Cys Gly
1 5 10 15

<210> 190
<211> 16
<212> PRT
<213> Homo sapiens

<400> 190
Glu Val Arg Gln Val Thr Cys Ile Arg Cys Arg Arg Gly Phe Leu Leu
1 5 10 15

<210> 191
<211> 15
<212> PRT
<213> Homo sapiens

<400> 191
Gly Gly Gly Gly Met Trp Glu Ala Trp Ser Cys Tyr Ala Cys Gly
1 5 10 15

<210> 192
<211> 16
<212> PRT
<213> Homo sapiens

<400> 192
Gly Trp Arg Trp Cys Gly Arg Cys Gly Ala Leu Trp Trp Arg Arg Val
1 5 10 15

<210> 193
<211> 485
<212> DNA
<213> Homo sapiens

<400> 193
aagcttgcca ccatggagac agacacactc ctgctatggg tactgctgct ctgggttcca 60
ggttccactg gtgacggatc catgagcgat aaaattattc acctgactga cgacagtttt 120
gacacggatg tactcaaagc ggacggggcg atcctcgtcg atttctgggc agagtggtgc 180
ggtccgaatt ccgtggttct gtgttcgcgt tgtgggcgtt tgtggcggtg gtcgtgtggg 240
actagtggtc cgtgcaaaat gatcgccccg attctggatg aaatcgctga cgaatatcag 300
ggcaaactga ccgttgcaaa actgaacatc gatcaaaacc ctggcactgc gccgaaatat 360
ggcatccgtg gtatcccgac tctgctgctg ttcaaaaacg gtgaagtggc ggcaaccaaa 420
gtgggtgcac tgtctaaagg tcagttgaaa gagttcctcg acgctaacct ggcgtaagcg 480
gccgc 485

<210> 194
<211> 485
<212> DNA
<213> Homo sapiens

<400> 194
aagcttgcca ccatggagac agacacactc ctgctatggg tactgctgct ctgggttcca 60
ggttccactg gtgacggatc catgagcgat aaaattattc acctgactga cgacagtttt 120
gacacggatg tactcaaagc ggacggggcg atcctcgtcg atttctgggc agagtggtgc 180
ggtccgaatt ccgggtggcg gtggtgtggt cggtgtgggg ctttgtggtg gcggcgtgtt 240
actagtggtc cgtgcaaaat gatcgccccg attctggatg aaatcgctga cgaatatcag 300
ggcaaactga ccgttgcaaa actgaacatc gatcaaaacc ctggcactgc gccgaaatat 360
ggcatccgtg gtatcccgac tctgctgctg ttcaaaaacg gtgaagtggc ggcaaccaaa 420
gtgggtgcac tgtctaaagg tcagttgaaa gagttcctcg acgctaacct ggcgtaagcg 480
gccgc 485

<210> 195
<211> 485
<212> DNA
<213> Homo sapiens

<400> 195
aagcttgcca ccatggagac agacacactc ctgctatggg tactgctgct ctgggttcca 60
ggttccactg gtgacggatc catgagcgat aaaattattc acctgactga cgacagtttt 120
gacacggatg tactcaaagc ggacggggcg atcctcgtcg atttctgggc agagtggtgc 180
ggtccgaatt ccgaggtgcg gcaggttacg tgtattaggt gtcgtcgggg ttttctgttg 240
actagtggtc cgtgcaaaat gatcgccccg attctggatg aaatcgctga cgaatatcag 300
ggcaaactga ccgttgcaaa actgaacatc gatcaaaacc ctggcactgc gccgaaatat 360
ggcatccgtg gtatcccgac tctgctgctg ttcaaaaacg gtgaagtggc ggcaaccaaa 420
gtgggtgcac tgtctaaagg tcagttgaaa gagttcctcg acgctaacct ggcgtaagcg 480
gccgc 485

<210> 196
<211> 485
<212> DNA
<213> Homo sapiens

<400> 196
aagcttgcca ccatggagac agacacactc ctgctatggg tactgctgct ctgggttcca 60
ggttccactg gtgacggatc catgagcgat aaaattattc acctgactga cgacagtttt 120
gacacggatg tactcaaagc ggacggggcg atcctcgtcg atttctgggc agagtggtgc 180
ggtccgaatt ccggtggtgg ggggatgatt tgggaggctt ggagttgtta tgcgtgtggg 240
actagtggtc cgtgcaaaat gatcgccccg attctggatg aaatcgctga cgaatatcag 300
ggcaaactga ccgttgcaaa actgaacatc gatcaaaacc ctggcactgc gccgaaatat 360
ggcatccgtg gtatcccgac tctgctgctg ttcaaaaacg gtgaagtggc ggcaaccaaa 420

gtgggtgcac tgtctaaagg tcagttgaaa gagttcctcg acgctaacct ggcgtaagcg 480
gccgc 485

<210> 197
<211> 485
<212> DNA
<213> Homo sapiens

<400> 197
aagcttgcca ccatggagac agacacactc ctgctatggg tactgctgct ctgggttcca 60
ggttccactg gtgacggatc catgagcgat aaaattattc acctgactga cgacagtttt 120
gacacggatg tactcaaagc ggacggggcg atcctcgtcg atttctgggc agagtggtgc 180
ggtccgaatt ccttgtggat tgggccgggt gatcagggtc tgtttcggcg ttttgttttt 240
actagtggtc cgtgcaaaat gatcgccccg attctggatg aaatcgctga cgaatatcag 300
ggcaaactga ccgttgcaaa actgaacatc gatcaaaacc ctggcactgc gccgaaatat 360
ggcatccgtg gtatcccgac tctgctgctg ttcaaaaacg gtgaagtggc ggcaaccaaa 420
gtgggtgcac tgtctaaagg tcagttgaaa gagttcctcg acgctaacct ggcgtaagcg 480
gccgc 485

<210> 198
<211> 476
<212> DNA
<213> Homo sapiens

<400> 198
aagcttgcca ccatggagac agacacactc ctgctatggg tactgctgct ctgggttcca 60
ggttccactg gtgacggatc cgtgtcttct gatcaaaatc atttccgagg agaaattgag 120
gaaaccatca ctgaaagctt tggtaatgat catagcacct tggatgggta ttccagaaga 180
accaccttgt cttcaaaaat gtatcacacc aaaggacaag aaggttctgt ttgtctccgg 240
tcatcagact gtgcctcagg attgtgttgt gctagacact tctggtccaa gatctgtaaa 300
cctgtcctga aagaaggtca agtgtgtacc aagcatagga gaaaaggctc tcatggacta 360
gaaatattcc agcgttgtta ctgtggagaa ggtctgtctt gccggataca gaaagatcac 420
catcaagcca gtaattcttc taggcttcac acttgtcaga gacactaagc ggccgc 476

<210> 199
<211> 539
<212> DNA
<213> Homo sapiens

<400> 199
aagcttgcca ccatggagac agacacactc ctgctatggg tactgctgct ctgggttcca 60
ggttccactg gtgacggatc ctgcgctagt cccacccgcg gaggggacgc gggcgtgcaa 120
atctgtctcg cctgcaggaa gcgccgaaaa cgctgcatgc gtcacgctat gtgctgcccc 180
gggaattact gcaaaaatgg aatatgtgtg tcttctgatc aaaatcattt ccgaggagaa 240
attgaggaaa ccatcactga aagctttggt aatgatcata gcaccttgga tgggtattcc 300
agaagaacca ccttgtcttc aaaaatgtat cacaccaaag gacaagaagg ttctgtttgt 360
ctccggtcat cagactgtgc ctcaggattg tgttgtgcta gacacttctg gtccaagatc 420
tgtaaacctg tcctgaaaga aggtcaagtg tgtaccaagc ataggagaaa aggctctcat 480
ggactagaaa tattccagcg ttgttactgt ggagaaggtc tgtcttgcta agcggccgc 539

<210> 200
<211> 416
<212> DNA
<213> Homo sapiens

<400> 200
aagcttgcca ccatgggcga taaaattatt cacctgactg acgacagttt tgacacggat 60

gtactcaaag cggacggggc gatcctcgtc gatttctggg cagagtggtg cggtccgaat 120
tcctatgcgt ggttgttttc ttgtagtagg tgtaggtggt ggttgccttg gactagtggt 180
ccgtgcaaaa tgatcgcccc gattctggat gaaatcgctg acgaatatca gggcaaactg 240
accgttgcaa aactgaacat cgatcaaaac cctggcactg cgccgaaata tggcatccgt 300
ggtatcccga ctctgctgct gttcaaaaac ggtgaagtgg cggcaaccaa agtgggtgca 360
ctgtctaaag gtcagttgaa agagttcctc gacgctaacc tggcgtaagc ggccgc 416

<210> 201
<211> 416
<212> DNA
<213> Homo sapiens

<400> 201
aagcttgcca ccatgggcga taaaattatt cacctgactg acgacagttt tgacacggat 60
gtactcaaac gccacggggc gatcctcgtc gatttctggg cagagtggtg cggtccgaat 120
tccatttgtg aggttgtgag gttgtggagt cggtatcctt ggtcttgggt gactagtggt 180
ccgtgcaaaa tgatcgcccc gattctggat gaaatcgctg acgaatatca gggcaaactg 240
accgttgcaa aactgaacat cgatcaaaac cctggcactg cgccgaaata tggcatccgt 300
ggtatcccga ctctgctgct gttcaaaaac ggtgaagtgg cggcaaccaa agtgggtgca 360
ctgtctaaag gtcagttgaa agagttcctc gacgctaacc tggcgtaagc ggccgc 416

<210> 202
<211> 422
<212> DNA
<213> Homo sapiens

<400> 202
aagcttgcca ccatgggcga taaaattatt cacctgactg acgacagttt tgacacggat 60
gtactcaaag cggacggggc gatcctcgtc gatttctggg cagagtggtg cggtccgaat 120
tccggttgta ctagtgcggt gtgtggtgct tgggctgagg cgggtaggtt ttattgtact 180
agtggtccgt gcaaaatgat cgccccgatt ctggatgaaa tcgctgacga atatcagggc 240
aaactgaccg ttgcaaaact gaacatcgat caaaaccctg gcactgcgcc gaaatatggc 300
atccgtggta tcccgactct gctgctgttc aaaaacggtg aagtggcggc aaccaaagtg 360
ggtgcactgt ctaaaggtca gttgaaagag ttcctcgacg ctaacctggc gtaagcggcc 420
gc 422

<210> 203
<211> 416
<212> DNA
<213> Homo sapiens

<400> 203
aagcttgcca ccatgggcga taaaattatt cacctgactg acgacagttt tgacacggat 60
gtactcaaag cggacggggc gatcctcgtc gatttctggg cagagtggtg cggtccgaat 120
tccttgtgga ttgggccggg tgatcagggt ctgtttcggc gttttgtttt tactagtggt 180
ccgtgcaaaa tgatcgcccc gattctggat gaaatcgctg acgaatatca gggcaaactg 240
accgttgcaa aactgaacat cgatcaaaac cctggcactg cgccgaaata tggcatccgt 300
ggtatcccga ctctgctgct gttcaaaaac ggtgaagtgg cggcaaccaa agtgggtgca 360
ctgtctaaag gtcagttgaa agagttcctc gacgctaacc tggcgtaagc ggccgc 416

<210> 204
<211> 154
<212> PRT
<213> Homo sapiens

<400> 204

```
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1               5                   10                  15
Gly Ser Thr Gly Asp Gly Ser Met Ser Asp Lys Ile Ile His Leu Thr
            20                  25                  30
Asp Asp Ser Phe Asp Thr Asp Val Leu Lys Ala Asp Gly Ala Ile Leu
        35                  40                  45
Val Asp Phe Trp Ala Glu Trp Cys Gly Pro Asn Ser Val Val Leu Cys
    50                  55                  60
Ser Arg Cys Gly Arg Leu Trp Arg Trp Ser Cys Gly Thr Ser Gly Pro
65                  70                  75                  80
Cys Lys Met Ile Ala Pro Ile Leu Asp Glu Ile Ala Asp Glu Tyr Gln
                85                  90                  95
Gly Lys Leu Thr Val Ala Lys Leu Asn Ile Asp Gln Asn Pro Gly Thr
            100                 105                 110
Ala Pro Lys Tyr Gly Ile Arg Gly Ile Pro Thr Leu Leu Leu Phe Lys
        115                 120                 125
Asn Gly Glu Val Ala Ala Thr Lys Val Gly Ala Leu Ser Lys Gly Gln
    130                 135                 140
Leu Lys Glu Phe Leu Asp Ala Asn Leu Ala
145                 150
```

<210> 205
<211> 154
<212> PRT
<213> Homo sapiens

<400> 205

```
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1               5                   10                  15
Gly Ser Thr Gly Asp Gly Ser Met Ser Asp Lys Ile Ile His Leu Thr
            20                  25                  30
Asp Asp Ser Phe Asp Thr Asp Val Leu Lys Ala Asp Gly Ala Ile Leu
        35                  40                  45
Val Asp Phe Trp Ala Glu Trp Cys Gly Pro Asn Ser Gly Trp Arg Trp
    50                  55                  60
Cys Gly Arg Cys Gly Ala Leu Trp Trp Arg Arg Val Thr Ser Gly Pro
65                  70                  75                  80
Cys Lys Met Ile Ala Pro Ile Leu Asp Glu Ile Ala Asp Glu Tyr Gln
                85                  90                  95
Gly Lys Leu Thr Val Ala Lys Leu Asn Ile Asp Gln Asn Pro Gly Thr
            100                 105                 110
Ala Pro Lys Tyr Gly Ile Arg Gly Ile Pro Thr Leu Leu Leu Phe Lys
        115                 120                 125
Asn Gly Glu Val Ala Ala Thr Lys Val Gly Ala Leu Ser Lys Gly Gln
    130                 135                 140
Leu Lys Glu Phe Leu Asp Ala Asn Leu Ala
145                 150
```

<210> 206
<211> 154
<212> PRT
<213> Homo sapiens

<400> 206

```
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1               5                   10                  15
```

Gly Ser Thr Gly Asp Gly Ser Met Ser Asp Lys Ile Ile His Leu Thr
20 25 30
Asp Asp Ser Phe Asp Thr Asp Val Leu Lys Ala Asp Gly Ala Ile Leu
35 40 45
Val Asp Phe Trp Ala Glu Trp Cys Gly Pro Asn Ser Glu Val Arg Gln
50 55 60
Val Thr Cys Ile Arg Cys Arg Arg Gly Phe Leu Leu Thr Ser Gly Pro
65 70 75 80
Cys Lys Met Ile Ala Pro Ile Leu Asp Glu Ile Ala Asp Glu Tyr Gln
85 90 95
Gly Lys Leu Thr Val Ala Lys Leu Asn Ile Asp Gln Asn Pro Gly Thr
100 105 110
Ala Pro Lys Tyr Gly Ile Arg Gly Ile Pro Thr Leu Leu Leu Phe Lys
115 120 125
Asn Gly Glu Val Ala Ala Thr Lys Val Gly Ala Leu Ser Lys Gly Gln
130 135 140
Leu Lys Glu Phe Leu Asp Ala Asn Leu Ala
145 150

<210> 207
<211> 154
<212> PRT
<213> Homo sapiens

<400> 207
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1 5 10 15
Gly Ser Thr Gly Asp Gly Ser Met Ser Asp Lys Ile Ile His Leu Thr
20 25 30
Asp Asp Ser Phe Asp Thr Asp Val Leu Lys Ala Asp Gly Ala Ile Leu
35 40 45
Val Asp Phe Trp Ala Glu Trp Cys Gly Pro Asn Ser Gly Gly Gly Gly
50 55 60
Met Ile Trp Glu Ala Trp Ser Cys Tyr Ala Cys Gly Thr Ser Gly Pro
65 70 75 80
Cys Lys Met Ile Ala Pro Ile Leu Asp Glu Ile Ala Asp Glu Tyr Gln
85 90 95
Gly Lys Leu Thr Val Ala Lys Leu Asn Ile Asp Gln Asn Pro Gly Thr
100 105 110
Ala Pro Lys Tyr Gly Ile Arg Gly Ile Pro Thr Leu Leu Leu Phe Lys
115 120 125
Asn Gly Glu Val Ala Ala Thr Lys Val Gly Ala Leu Ser Lys Gly Gln
130 135 140
Leu Lys Glu Phe Leu Asp Ala Asn Leu Ala
145 150

<210> 208
<211> 154
<212> PRT
<213> Homo sapiens

<400> 208
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1 5 10 15
Gly Ser Thr Gly Asp Gly Ser Met Ser Asp Lys Ile Ile His Leu Thr

```
                20                  25                  30
Asp Asp Ser Phe Asp Thr Asp Val Leu Lys Ala Asp Gly Ala Ile Leu
        35                  40                  45
Val Asp Phe Trp Ala Glu Trp Cys Gly Pro Asn Ser Leu Trp Ile Gly
    50                  55                  60
Pro Gly Asp Gln Gly Leu Phe Arg Arg Phe Val Phe Thr Ser Gly Pro
65                  70                  75                  80
Cys Lys Met Ile Ala Pro Ile Leu Asp Glu Ile Ala Asp Glu Tyr Gln
                85                  90                  95
Gly Lys Leu Thr Val Ala Lys Leu Asn Ile Asp Gln Asn Pro Gly Thr
            100                 105                 110
Ala Pro Lys Tyr Gly Ile Arg Gly Ile Pro Thr Leu Leu Leu Phe Lys
        115                 120                 125
Asn Gly Glu Val Ala Ala Thr Lys Val Gly Ala Leu Ser Lys Gly Gln
    130                 135                 140
Leu Lys Glu Phe Leu Asp Ala Asn Leu Ala
145                 150
```

<210> 209
<211> 151
<212> PRT
<213> Homo sapiens

<400> 209

```
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1               5                   10                  15
Gly Ser Thr Gly Asp Gly Ser Val Ser Ser Asp Gln Asn His Phe Arg
            20                  25                  30
Gly Glu Ile Glu Glu Thr Ile Thr Glu Ser Phe Gly Asn Asp His Ser
        35                  40                  45
Thr Leu Asp Gly Tyr Ser Arg Arg Thr Thr Leu Ser Ser Lys Met Tyr
    50                  55                  60
His Thr Lys Gly Gln Glu Gly Ser Val Cys Leu Arg Ser Ser Asp Cys
65                  70                  75                  80
Ala Ser Gly Leu Cys Cys Ala Arg His Phe Trp Ser Lys Ile Cys Lys
                85                  90                  95
Pro Val Leu Lys Glu Gly Gln Val Cys Thr Lys His Arg Arg Lys Gly
            100                 105                 110
Ser His Gly Leu Glu Ile Phe Gln Arg Cys Tyr Cys Gly Glu Gly Leu
        115                 120                 125
Ser Cys Arg Ile Gln Lys Asp His His Gln Ala Ser Asn Ser Ser Arg
    130                 135                 140
Leu His Thr Cys Gln Arg His
145                 150
```

<210> 210
<211> 172
<212> PRT
<213> Homo sapiens

<400> 210

```
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1               5                   10                  15
Gly Ser Thr Gly Asp Gly Ser Cys Ala Ser Pro Thr Arg Gly Gly Asp
            20                  25                  30
Ala Gly Val Gln Ile Cys Leu Ala Cys Arg Lys Arg Arg Lys Arg Cys
```

35 40 45
Met Arg His Ala Met Cys Cys Pro Gly Asn Tyr Cys Lys Asn Gly Ile
50 55 60
Cys Val Ser Ser Asp Gln Asn His Phe Arg Gly Glu Ile Glu Glu Thr
65 70 75 80
Ile Thr Glu Ser Phe Gly Asn Asp His Ser Thr Leu Asp Gly Tyr Ser
85 90 95
Arg Arg Thr Thr Leu Ser Ser Lys Met Tyr His Thr Lys Gly Gln Glu
100 105 110
Gly Ser Val Cys Leu Arg Ser Ser Asp Cys Ala Ser Gly Leu Cys Cys
115 120 125
Ala Arg His Phe Trp Ser Lys Ile Cys Lys Pro Val Leu Lys Glu Gly
130 135 140
Gln Val Cys Thr Lys His Arg Arg Lys Gly Ser His Gly Leu Glu Ile
145 150 155 160
Phe Gln Arg Cys Tyr Cys Gly Glu Gly Leu Ser Cys
165 170

<210> 211
<211> 131
<212> PRT
<213> Homo sapiens

<400> 211
Met Gly Asp Lys Ile Ile His Leu Thr Asp Asp Ser Phe Asp Thr Asp
1 5 10 15
Val Leu Lys Ala Asp Gly Ala Ile Leu Val Asp Phe Trp Ala Glu Trp
20 25 30
Cys Gly Pro Asn Ser Tyr Ala Trp Leu Phe Ser Cys Ser Arg Cys Arg 35
40 45
Trp Trp Leu Pro Trp Thr Ser Gly Pro Cys Lys Met Ile Ala Pro Ile
50 55 60
Leu Asp Glu Ile Ala Asp Glu Tyr Gln Gly Lys Leu Thr Val Ala Lys
65 70 75 80
Leu Asn Ile Asp Gln Asn Pro Gly Thr Ala Pro Lys Tyr Gly Ile Arg
85 90 95
Gly Ile Pro Thr Leu Leu Leu Phe Lys Asn Gly Glu Val Ala Ala Thr
100 105 110
Lys Val Gly Ala Leu Ser Lys Gly Gln Leu Lys Glu Phe Leu Asp Ala
115 120 125
Asn Leu Ala
130

<210> 212
<211> 131
<212> PRT
<213> Homo sapiens

<400> 212
Met Gly Asp Lys Ile Ile His Leu Thr Asp Asp Ser Phe Asp Thr Asp
1 5 10 15
Val Leu Lys Arg His Gly Ala Ile Leu Val Asp Phe Trp Ala Glu Trp
20 25 30
Cys Gly Pro Asn Ser Ile Cys Glu Val Val Arg Leu Trp Ser Arg Tyr
35 40 45
Pro Trp Ser Trp Val Thr Ser Gly Pro Cys Lys Met Ile Ala Pro Ile

```
    50                  55                  60
Leu Asp Glu Ile Ala Asp Glu Tyr Gln Gly Lys Leu Thr Val Ala Lys
65                  70                  75                  80
Leu Asn Ile Asp Gln Asn Pro Gly Thr Ala Pro Lys Tyr Gly Ile Arg
                85                  90                  95
Gly Ile Pro Thr Leu Leu Leu Phe Lys Asn Gly Glu Val Ala Ala Thr
            100                 105                 110
Lys Val Gly Ala Leu Ser Lys Gly Gln Leu Lys Glu Phe Leu Asp Ala
        115                 120                 125
Asn Leu Ala
    130
```

<210> 213
<211> 133
<212> PRT
<213> Homo sapiens

<400> 213

```
Met Gly Asp Lys Ile Ile His Leu Thr Asp Asp Ser Phe Asp Thr Asp
1               5                   10                  15
Val Leu Lys Ala Asp Gly Ala Ile Leu Val Asp Phe Trp Ala Glu Trp
            20                  25                  30
Cys Gly Pro Asn Ser Gly Cys Thr Ser Ala Val Cys Gly Ala Trp Ala
        35                  40                  45
Glu Ala Gly Arg Phe Tyr Cys Thr Ser Gly Pro Cys Lys Met Ile Ala
    50                  55                  60
Pro Ile Leu Asp Glu Ile Ala Asp Glu Tyr Gln Gly Lys Leu Thr Val
65                  70                  75                  80
Ala Lys Leu Asn Ile Asp Gln Asn Pro Gly Thr Ala Pro Lys Tyr Gly
                85                  90                  95
Ile Arg Gly Ile Pro Thr Leu Leu Leu Phe Lys Asn Gly Glu Val Ala
            100                 105                 110
Ala Thr Lys Val Gly Ala Leu Ser Lys Gly Gln Leu Lys Glu Phe Leu
        115                 120                 125
Asp Ala Asn Leu Ala
    130
```

<210> 214
<211> 131
<212> PRT
<213> Homo sapiens

<400> 214

```
Met Gly Asp Lys Ile Ile His Leu Thr Asp Asp Ser Phe Asp Thr Asp
1               5                   10                  15
Val Leu Lys Ala Asp Gly Ala Ile Leu Val Asp Phe Trp Ala Glu Trp
            20                  25                  30
Cys Gly Pro Asn Ser Leu Trp Ile Gly Pro Gly Asp Gln Gly Leu Phe
        35                  40                  45
Arg Arg Phe Val Phe Thr Ser Gly Pro Cys Lys Met Ile Ala Pro Ile
    50                  55                  60
Leu Asp Glu Ile Ala Asp Glu Tyr Gln Gly Lys Leu Thr Val Ala Lys
65                  70                  75                  80
Leu Asn Ile Asp Gln Asn Pro Gly Thr Ala Pro Lys Tyr Gly Ile Arg
                85                  90                  95
```

```
Gly Ile Pro Thr Leu Leu Leu Phe Lys Asn Gly Glu Val Ala Ala Thr
            100                 105                 110
Lys Val Gly Ala Leu Ser Lys Gly Gln Leu Lys Glu Phe Leu Asp Ala
        115                 120                 125
Asn Leu Ala
    130
```

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 November 2002 (21.11.2002)

PCT

(10) International Publication Number
WO 02/092015 A3

(51) International Patent Classification[7]: A61K 39/395, 39/00, 39/38

(21) International Application Number: PCT/US02/15982

(22) International Filing Date: 17 May 2002 (17.05.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/291,311	17 May 2001 (17.05.2001)	US
60/353,058	1 February 2002 (01.02.2002)	US
60/361,293	4 March 2002 (04.03.2002)	US

(71) Applicants *(for all designated States except US)*: **GENOME THERAPEUTICS CORPORATION** [US/US]; 100 Beaver Street, Waltham, MA 02453 (US). **WYETH** [US/US]; Five Giralda Farms, Madison, NJ 07928 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ALLEN, Kristina** [US/US]; 11 Oliver Lane, Hopkinton, MA 01748-3108 (US). **ANISOWICZ, Anthony** [US/US]; 50 Upham Street, West Newton, MA 02465 (US). **BHAT, Bheem, M.** [IN/US]; 1214 Mayapple Lane, West Chester, PA 19380 (US). **DAMAGNEZ, Veronique** [FR/US]; 125 Water Street, Framingham, MA 01701 (US). **ROBINSON, John, Allen** [US/US]; 23 Webb Road, Downingtown, PA 19335 (US). **YAWORSKY, Paul, J.** [US/US]; 13 Hobart Lane, Rockland, MA 02370 (US).

(74) Agents: **REA, Teresa, Stanek** et al.; Burns, Doane, Swecker & Mathis L.L.P., P.O. Box 1404, Alexandria, VA 22313-1404 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report: 23 October 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/092015 A3

(54) Title: REAGENTS AND METHODS FOR MODULATING DKK-MEDIATED INTERACTIONS

(57) Abstract: The present invention provides reagents, compounds, compositions, and methods relating to novel interactions of the extracellular domain of LRP5, HBM (a variant of LRP5), and/or LRP6 with Dkk, including Dkk-1. The various nucleic acids, polypeptides, antibodies, assay methods, diagnostic methods, and methods of treatment of the present invention are related to and impact on Dkk, LRP5, LRP6, HBM, and Wnt signaling. Dkk, LRP5, LRP6, HBM, and Wnt are implicated in bone and lipid cellular signaling. Thus, the present invention provides reagents and methods for modulating lipid levels and/or bone mass and is useful in the treatment and diagnosis of abnormal lipid levels and bone mass disorders, such as osteoporosis.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/15982

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : A61K 39/395, 00, 38

US CL : 424/130.1, 184.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 424/130.1, 184.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	GONG,Y et al. LDL RECEPTOR-RELATED PROTEIN 5 (LRP-5) AFFECTS BONE ACCRUAL AND EYE DEVELOPMENT, CELL , 2001 VOL.107, PAGES 513-523	1-11 and 13-23
	ZORN A. WNT SIGNALLING: ANTABONISTIC DICKKOPFS, CURRENT BIOLOGY, 2001, VOL.11, No.15, PAGES R592-R595	1-11 and 13-23
A	WO 9846743 A1 (THE WELL-COME TRUST LIMITED), 22 OCTOBER 1998 (22.10.98) see entire document.	1-11 and 13-23

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 March 2003 (18.03.2003)	08 AUG 2003

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Valerie Bell-Harris Michail A Belyavskyi Telephone No. 703/308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US02/15982

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claim Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-11 and 13-23

Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

I. Claims 1-11 and 13-23, drawn to a method of regulating LRP5 activity in a subject, comprising administering a composition which inhibits Dkk binding to LRP5.

II. Claims 1-11 and 13-20, drawn to a method of regulating LRP5 activity in a subject, comprising administering a composition which inhibits Dkk binding to LRP6.

III. Claims 1-11 and 13-20, drawn to a method of regulating LRP5 activity in a subject, comprising administering a composition which inhibits Dkk binding to HBM.

IV. Claims 1-10, 12-14 and 16-23 drawn to a method of regulating LRP5 activity in a subject, comprising administering a composition which enhances Dkk binding to LRP5.

V. Claims 1-10, 12-14 and 16-20, drawn to a method of regulating LRP5 activity in a subject, comprising administering a composition which enhances Dkk binding to LRP6.

VI. Claims 1-10, 12-14 and 16-20, drawn to a method of regulating LRP5 activity in a subject, comprising administering a composition which enhances Dkk binding to HBM.

VII. Claims 2 - 4, 7-11, 13-20, 21-27, drawn to a method of regulating Dkk-Wnt pathway activity, comprising administering a composition which modulates Dkk activity, or interaction with LRP5, wherein said composition inhibits Dkk binding to LRP5.

VIII. Claims 2 - 4, 7-11, 13-20, 24-27, drawn to a method of regulating Dkk-Wnt pathway activity, comprising administering a composition which modulates Dkk activity, or interaction with LRP6, wherein said composition inhibits Dkk binding to LRP6.

IX. Claims 2 - 4, 7-11, 13-20, 24-27, drawn to a method of regulating Dkk-Wnt pathway activity, comprising administering a composition which modulates Dkk activity, or interaction with HBM, wherein said composition inhibits Dkk binding to HBM.

X. Claims 2 - 4, 7-10, 12-27, drawn to a method of regulating Dkk-Wnt pathway activity, comprising administering a composition which modulates Dkk activity, or interaction with RP5, wherein said composition enhances Dkk binding to LRP5.

XI. Claims 2 - 4, 7-10, 12-20, 24-27, drawn to a method of regulating Dkk-Wnt pathway activity, comprising administering a composition which modulates Dkk activity, or interaction with LRP6, wherein said composition enhances Dkk binding to LRP6.

XII. Claims 2 - 4, 7-10, 12-20, 24-27, drawn to a method of regulating Dkk-Wnt pathway activity, comprising administering a composition which modulates Dkk activity, or interaction with HBM, wherein said composition enhances Dkk binding to HBM.

XIII. Claims 2 - 4, 7-11, 13-23, 28-32 drawn to a method of modulating bone mass in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP5, wherein said composition inhibits Dkk binding to LRP5.

XIV. Claims 2 - 4, 7-11, 13-23, 28-32 drawn to a method of modulating bone mass in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP6, wherein said composition inhibits Dkk binding to LRP6.

XV. Claims 2 - 4, 7-11, 13-23, 28-32 drawn to a method of modulating bone mass in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with HBM, wherein said composition inhibits Dkk binding to HBM.

INTERNATIONAL SEARCH REPORT

XVI. Claims 2 - 4, 7-10, 12-23, 28-32 drawn to a method of modulating bone mass in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP5, wherein said composition enhances Dkk binding to LRP5.

XIV. Claims 2 - 4, 7-10, 12-20, 28-32 drawn to a method of modulating bone mass in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP6, wherein said composition enhances Dkk binding to LRP6.

XV. Claims 2 - 4, 7-10, 12-20, 28-32 drawn to a method of modulating bone mass in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with HBM, wherein said composition enhances Dkk binding to HBM.

XVI. Claims 2 - 4, 7-11, 13-23, 33-35 drawn to a method of modulating lipid levels in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP5, wherein said composition inhibits Dkk binding to LRP5.

XVII. XVI. Claims 2 - 4, 7-11, 13-23, 33-35 drawn to a method of modulating lipid levels in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP6, wherein said composition inhibits Dkk binding to LRP6.

XVIII. Claims 2 - 4, 7-11, 13-23, 33-35 drawn to a method of modulating lipid levels in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with HBM, wherein said composition inhibits Dkk binding to HBM.

XIX. Claims 2 - 4, 7-10, 12-23, 33-35 drawn to a method of modulating lipid levels in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP5, wherein said composition enhances Dkk binding to LRP5.

XX. Claims 2 - 4, 7-10, 12-20, 33-35 drawn to a method of modulating lipid levels in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP6, wherein said composition enhances Dkk binding to LRP6.

XXI. Claims 2 - 4, 7-10, 12-20, 33-35 drawn to a method of modulating lipid levels in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with HBM, wherein said composition enhances Dkk binding to HBM.

XXII. Claims 2, 35 and 36 drawn to a method of diagnosing low or high bone mass and /or high or low lipid levels in a subject comprising examining expression of Dkk.

XXIII. Claims 2, 35 and 36 drawn to a method of diagnosing low or high bone mass and /or high or low lipid levels in a subject comprising examining expression of LRP5.

XXIV. Claims 2, 35 and 36 drawn to a method of diagnosing low or high bone mass and /or high or low lipid levels in a subject comprising examining expression of LRP6.

XXV. Claims 2, 35 and 36 drawn to a method of diagnosing low or high bone mass and /or high or low lipid levels in a subject comprising examining expression of HBM.

XXVI. Claims 2, 37-43, 44-47, drawn to a method of screening for a compound which modulates the interaction of DKK with LRP5.

XXVII. Claims 2, 37-43, 47, drawn to a method of screening for a compound which modulates the interaction of DKK with LRP6.

XXVIII. Claims 2, 37-43, 47, drawn to a method of screening for a compound which modulates the interaction of DKK with HBM.

XXIX. Claims 2,41-43, 48-49, drawn to a method of screening for a compound which modulates the interaction of Dkk with a Dkk interacting proteins.

XXX. Claims 50-63 drawn to a composition comprising a LRP5 and a pharmaceutical acceptable carrier thereof.

XXXI. Claims 50- 59, 63 drawn to a composition comprising a LRP6 and a pharmaceutical acceptable carrier thereof.

XXXII. Claims 50- 59, 63 drawn to a composition comprising a HBM and a pharmaceutical acceptable carrier thereof.

XXXIII. Claims 2 and 64 drawn to a method for identifying compound which modulate Dkk and LRP5 interaction.

XXXIV. Claims 2 and 64 drawn to a method for identifying compound which modulate Dkk and LRP6 interaction.

XXXV. Claims 2 and 64 drawn to a method for identifying compound which modulate Dkk and HBM interaction.

XXXVI. Claims 2 and 65 drawn to a method of identifying binding partners for a Dkk protein.

XXXVII. Claims 66-68 drawn to a nucleic acid and a vector encoding a Dkk interacting protein.

XXXVIII. Claims 69-90, drawn to a method of detecting a modulatory activity of a compound, wherein the first peptide is a Dkk peptide and the second peptide is a LRP5.

XXXIX. Claims 69-87, drawn to a method of detecting a modulatory activity of a compound, wherein the first peptide is a Dkk peptide and the second peptide is a LRP6.

XL. Claims 69-87, drawn to a method of detecting a modulatory activity of a compound, wherein the first peptide is a Dkk peptide and the second peptide is a HBM.

XLI. Claims 91-92 drawn to a transgenic animal.

XLII. Claims 2 and 93 drawn to a method for identifying potential compound which modulate Dkk activity..

XLIII - LXII. Claim 94, drawn to one specific peptide aptamer of one specific SEQ ID NOs : 171-88; 189-192.

LXIII- LXXIX. Claims 95-97, drawn to an antibody which specifically recognizes and binds to specific peptides of SEQ ID NOs : 110-127.

LXXX . Claims 2, 98-100 , drawn to a method of identifying Dkk interacting protein which modulate the interaction of Dkk with Wnt signaling pathway.

LXXXI. Claims 2, 25 and 101-104, drawn to a method for identifying Dkk interacting proteins.

LXXXII. Claims 2, 105-106, drawn to a method for identifying compounds which modulate Dkk and LRP5 interaction.

LXXXIII. Claims 2, 105-106, drawn to a method for identifying compounds which modulate Dkk and LRP6 interaction.

LXXXIV. Claims 2, 105-106, drawn to a method for identifying compounds which modulate Dkk and HBM interaction.

LXXXV. Claims 2, 25, 107-110, drawn to a method for identifying compound which modulate the interaction of Dkk with Wnt signaling pathway.

LXXXVI. Claims 2, 111, drawn to a method of testing compounds that modulate Dkk-mediated activity in a mammal.

LXXXVII. Claims 2, 112, 113, drawn to method of screening for compound or composition which modulate the interaction of Dkk and Dkk interacting protein.

LXXXVIII-CIX. Claim 114 drawn to antibody which recognizes and binds to one specific SEQ ID NOs: 171-192.

The inventions listed as Groups 1-109 do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The invention listed as groups 1-109 do not related to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, hey lack the same or corresponding special technical features for the following reasons:

The special technical features of Group I is considered a method of regulating LRP5 activity in a subject, comprising administering a composition which inhibits Dkk binding to LRP5.

The special technical features of Group II is considered a method of regulating LRP5 activity in a subject, comprising administering a composition which inhibits Dkk binding to LRP6.

The special technical features of Group III is considered a method of regulating LRP5 activity in a subject, comprising administering a composition which inhibits Dkk binding to HBM.

The special technical features of Group IV is considered a method of regulating LRP5 activity in a subject, comprising administering a composition which enhances Dkk binding to LRP5.

The special technical features of Group V is considered a method of regulating LRP5 activity in a subject, comprising administering a composition which enhances Dkk binding to LRP6.

The special technical features of Group VI is considered a method of regulating LRP5 activity in a subject, comprising administering a composition which enhances Dkk binding to HBM.

INTERNATIONAL SEARCH REPORT

The special technical features of Group VII is considered a method of regulating Dkk-Wnt pathway activity, comprising administering a composition which modulates Dkk activity, or interaction with LRP5, wherein said composition inhibits Dkk binding to LRP5.

The special technical features of Group VIII is considered a method of regulating Dkk-Wnt pathway activity, comprising administering a composition which modulates Dkk activity, or interaction with LRP6, wherein said composition inhibits Dkk binding to LRP6.

The special technical features of Group IX is considered a method of regulating Dkk-Wnt pathway activity, comprising administering a composition which modulates Dkk activity, or interaction with HBM, wherein said composition inhibits Dkk binding to HBM.

The special technical features of Group X is considered a method of regulating Dkk-Wnt pathway activity, comprising administering a composition which modulates Dkk activity, or interaction with RP5, wherein said composition enhances Dkk binding to LRP5.

The special technical features of Group XI is considered a method of regulating Dkk-Wnt pathway activity, comprising administering a composition which modulates Dkk activity, or interaction with LRP6, wherein said composition enhances Dkk binding to LRP6.

The special technical features of Group XII is considered a method of regulating Dkk-Wnt pathway activity, comprising administering a composition which modulates Dkk activity, or interaction with HBM, wherein said composition enhances Dkk binding to HBM.

The special technical features of Group XIII is considered a method of modulating bone mass in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP5, wherein said composition inhibits Dkk binding to LRP5.

The special technical features of Group XIV is considered a method of modulating bone mass in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP6, wherein said composition inhibits Dkk binding to LRP6.

The special technical features of Group XV is considered a method of modulating bone mass in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with HBM, wherein said composition inhibits Dkk binding to HBM.

The special technical features of Group XVI is considered a method of modulating bone mass in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP5, wherein said composition enhances Dkk binding to LRP5.

The special technical features of Group XIV is considered a method of modulating bone mass in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP6, wherein said composition enhances Dkk binding to LRP6.

The special technical features of Group XV is considered a method of modulating bone mass in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with HBM, wherein said composition enhances Dkk binding to HBM.

The special technical features of Group XVI is considered a method of modulating lipid levels in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP5, wherein said composition inhibits Dkk binding to LRP5.

The special technical features of Group XVII is considered a method of modulating lipid levels in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP6, wherein said composition inhibits Dkk binding to LRP6.

The special technical features of Group XVIII is considered a method of modulating lipid levels in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with HBM, wherein said composition inhibits Dkk binding to HBM.

The special technical features of Group XIX is considered a method of modulating lipid levels in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP5, wherein said composition enhances Dkk binding to LRP5.

The special technical features of Group XX is considered a method of modulating lipid levels in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with LRP6, wherein said composition enhances Dkk binding to LRP6.

The special technical features of Group XXI is considered a method of modulating lipid levels in a subject comprising administering to the subject a composition which modulates DKK activity or interaction with HBM, wherein said composition enhances Dkk binding to HBM.

The special technical features of Group XXII is considered a method of diagnosing low or high bone mass and /or high or low lipid levels in a subject comprising examining expression of Dkk.

The special technical features of Group XXIII is considered a method of diagnosing low or high bone mass and /or high or low lipid levels in a subject comprising examining expression of LRP5.

The special technical features of Group XXIV is considered a method of diagnosing low or high bone mass and /or high or low lipid levels in a subject comprising examining expression of LRP6.

The special technical features of Group XXV is considered a method of diagnosing low or high bone mass and /or high or low lipid levels in a subject comprising examining expression of HBM.

The special technical features of Group XXVI is considered a method of screening for a compound which modulates the interaction of DKK with LRP5.

The special technical features of Group XXVII is considered a method of screening for a compound which modulates the interaction of DKK with LRP6.

The special technical features of Group XXVIII is considered a method of screening for a compound which modulates the interaction of DKK with HBM.

The special technical features of Group XXIX is considered a method of screening for a compound which modulates the interaction of Dkk with a Dkk interacting proteins.

The special technical features of Group XXX is considered a composition comprising a LRP5 and a pharmaceutical acceptable carrier thereof.

The special technical features of Group XXXI is considered a composition comprising a LRP6 and a pharmaceutical acceptable carrier thereof.

The special technical features of Group XXXII is considered a composition comprising a HBM and a pharmaceutical acceptable carrier thereof.

The special technical features of Group XXXIII is considered a method for identifying compound which modulate Dkk and LRP5 interaction.

The special technical features of Group XXXIV is considered a method for identifying compound which modulate Dkk and LRP6 interaction.

The special technical features of Group XXXV is considered a method for identifying compound which modulate Dkk and HBM interaction.

The special technical features of Group XXXVI is considered a method of identifying binding partners for a Dkk protein.

The special technical features of Group XXXVII is considered a nucleic acid and a vector encoding a Dkk interacting protein.

The special technical features of Group XXXVIII is considered a method of detecting a modulatory activity of a compound, wherein the first peptide is a Dkk peptide and the second peptide is a LRP5.

The special technical features of Group XXXIX is considered a method of detecting a modulatory activity of a compound, wherein the first peptide is a Dkk peptide and the second peptide is a LRP6.

The special technical features of Group XL is considered a method of detecting a modulatory activity of a compound, wherein the first peptide is a Dkk peptide and the second peptide is a HBM.

The special technical features of Group XLI is considered a transgenic animal.

The special technical features of Group XLII is considered a method for identifying potential compound which modulate Dkk activity.

The special technical features of Group XLIII - LXII is considered one specific peptide aptamer of one specific SEQ ID NOs : 171-88; 189-192.

The special technical features of Group LXIII- LXXIX is considered an antibody which specifically recognizes and binds to specific peptides of SEQ ID NOs : 110-127.

The special technical features of Group LXXX is considered a method of identifying Dkk interacting protein which modulate the interaction of Dkk with Wnt signaling pathway.

The special technical features of Group LXXXI is considered a method for identifying Dkk interacting proteins.

The special technical features of Group LXXXII is considered a method for identifying compounds which modulate Dkk and LRP5 interaction.

The special technical features of Group LXXXIII is considered a method for identifying compounds which modulate Dkk and LRP6 interaction.

The special technical features of Group LXXXIV is considered a method for identifying compounds which modulate Dkk and HBM interaction.

The special technical features of Group LXXXV is considered a method for identifying compound which modulate the interaction of Dkk with Wnt signaling pathway.

The special technical features of Group LXXXVI is considered a method of testing compounds that modulate Dkk-mediated activity in a mammal.

The special technical features of Group LXXXVII is considered a method of screening for compound or composition which modulate the interaction of Dkk and Dkk interacting protein.

The special technical features of Group LXXXVIII-CIX. is considered an antibody which recognizes and binds to one specific SEQ ID NOs: 171-192.

Accordingly, Groups I-CIX are not so linked by the same or corresponding special technical feature within meaning of PCT Rule 13.2 so as to form a single general inventive concept.

Continuation of B. FIELDS SEARCHED Item 3:

INTERNATIONAL SEARCH REPORT

PCT/US02/15982

Biosis, CAPLUS, SciSearch, Medline, EMBASE, WEST, USPATFULL, PCTFULL
search terms; Allen K; Anisowicz,A; Bhat,B; Damagnez,V, Robinson,J;Yaworsky,P; DKK, Dkk1, LRP5, SEQ ID NO:28, protein OST262; osteoporosis.